

# 強化学習による人型モデルの直立姿勢制御

電子情報研究部 堀 彰悟 (石川工業高等専門学校 3年)

## 研究の動機

2017年の"AlphaGoが囲碁で世界最強の囲碁棋士に勝利した"という報道から人工知能に興味を持った。そして、ネットで調べて強化学習という手法を知り、人や動物たこうしたらこうなるというパターンを経験の積み重ねで学習するのに似ているという特徴に強く惹かれた。高専でプログラミングや数学をある程度習得したので、実際に強化学習に取り組んでみてどのような結果が得られるのか経験を積みたいと思い、今回の課題取り組んだ。

## 強化学習とは

今回、用いた「強化学習」とは、

- (1) 与えられた環境と現在の状態に対して行動を決定し、その結果、次の状態に移動し、報酬を得る。
- (2) より多くの報酬を得られるよう行動の決定方針を修正する。

(1),(2)を繰り返すことで最も報酬が多く得られる行動方針を求める手法である。

## 研究方法：直立姿勢の学習

本研究では、Unity ML-Agents[1] を用いて、直立姿勢の学習を行った。図1のようにキャラクターモデルのボーンにそれぞれ剛体を割り当て、関節の向き(3次元ベクトル)を出力として学習させた。剛体の制御は出力ベクトルの位置にそのまま移動させるのではなく、出力ベクトルの向きになるような力(トルク)を加えている。

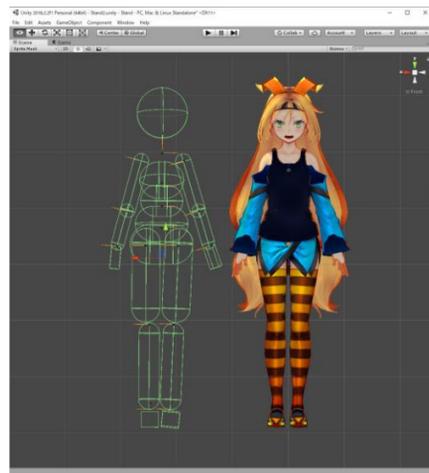


図1：キャラクターモデルとその剛体

自然な立ち方を学習させるため、報酬の与え方を適切に設定する必要がある。そのため、まず頭の位置が一定以上だったら報酬を与える設定で学習を行い、その学習結果から与え方を工夫して学習させる。これを何度か繰り返してより自然な立ち方に近づけていく。

### 学習モデルA：頭が一定以上だったら報酬を与える

図2に結果を示す。地面をけて後ろに下がっていくようになった。開始数秒はその場にとどまっているがだんだんと後ろに傾き、倒れないように膝を曲げ伸ばしするようになった。その場にとどまっていたので、モデルBでは頭の位置を報酬に加えることにした。

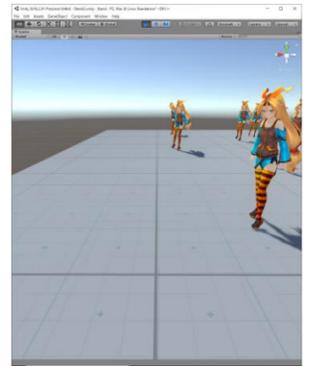


図2：モデルAの学習結果。時系列：左->中->右。後に移動している様子。

### 学習モデルB：Aの報酬-頭の位置のずれ

図3に結果を示す。多少最初の位置からずれるが安定的に立っていられたようになった。Aのモデルは大きくバランスを崩して倒れることがあったが、このモデルはほとんどバランスを崩さずかなり長い時間たっていた。ただ、足を肩幅に開いて仁王立ちしてしまうため、それを改善すべく次は足や背骨の地面との向きを報酬に加えることにした。

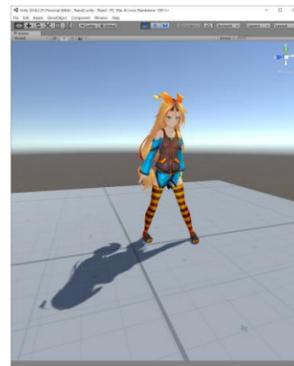


図3：モデルBの学習結果。仁王立ちしている様子

### 学習モデルC：Bの報酬+背骨・両足それぞれの上ベクトルと地面の法線との内積

図4に結果を示す。足をなるべく地面と垂直にして立つようになった。理想の形に近づいてきたが、バランスをとるための動きで足が細かく震える動作をしてしまう結果になった。最小限の動きで、動いてないように見えるように、立つことが理想のため、次はそれぞれのパーツの速度を報酬から減点することにした。



図4：モデルCの学習結果。足が地面と垂直に立つ様子



図5：モデルDの学習結果。前傾姿勢になった様子

### 学習モデルD：Cの報酬-各パーツの速度

図5に結果を示す。モデルCより動きは小さくなったが、前傾姿勢になってしまうという結果になった。徐々に後ろへ下がっていき、頭的位置を変えないように前傾姿勢になっていってしまうという結果になった。

各学習の経過を図6~13に示す。A,B,C,Dがそれぞれ赤、灰、水、緑色に該当する。繰り返し回数を増やすごとに、A,Bは評価値が増加しているが、C,Dは評価値が増加していない。この理由として、内積の値がエピソード長とあまり関係がないためだと考えられる

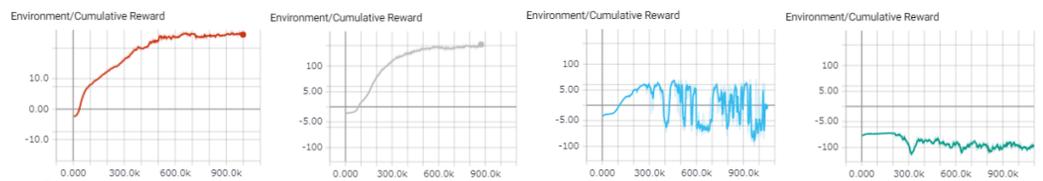


図6：Aの評価値の推移

図7：Bの評価値の推移

図8：Cの評価値の推移

図9：Dの評価値の推移

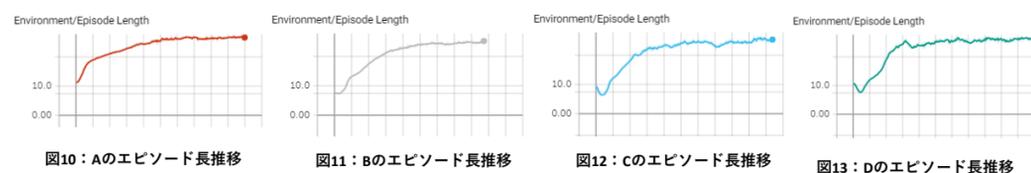


図10：Aのエピソード長推移

図11：Bのエピソード長推移

図12：Cのエピソード長推移

図13：Dのエピソード長推移

## 考察

今回の学習はA,B,C,Dそれぞれの結果について、ある程度目的に沿った形となったのでうまくいったといえる。

本研究の理想とする直立姿勢を学習させるにはもう少し複雑な報酬の与え方が必要だと思われる。あるいは、学習を足首だけにしてしまうというやり方も考えられる。

## 参考文献

- [1] 布留川 英一, "Unityではじめる機械学習・強化学習Unity ML-Agents実践ゲームプログラミング学習", ポーンデジタル, 2018