

# ペットロボットを用いた児童向け強化学習体験システム

大谷 紀子 研究室

2172043 姜 勝太

## 1. 背景と目的

プログラミング学習の支援方法の一つとして、クアッドクローラーの使用が挙げられる。クアッドクローラーとは、ハードウェアとソフトウェアの両側面から作る楽しさを味わうことができる、多関節四足歩行ロボットである。自身で組み込んだプログラムを実行すると、クアッドクローラーが動作し、出力結果を視覚的に確認することができる。

星[1]は、小学生がアルゴリズムに興味を持つことを目的として対話型の粒子群最適化（PSO : Particle Swarm Optimization）を用いた動作学習手法を提案した。星の手法では、ポーズ、ダンス、回転を学習対象の動作としている。評価実験の結果、アルゴリズムへの興味の向上に効果があることが示されたが、「もっと犬のような動きのロボットが良い」などの意見も得られた。

クアッドクローラーを犬のように動かすことで親しみやすさが増し、ペットとしての役割を持たせることができる。また、個人ではなく集団でペットとしてしつけ、感情的な相互作用を働かせることで、プログラミングへの興味がより深まると考えられる。

本研究では、小学生がプログラミングに興味を持つことを目的とし、クアッドクローラーがペットのような動作を強化学習アルゴリズムによって学習するシステムを構築する。

## 2. ユーザの操作手順

本システムではユーザがリモコン操作、および PC での報酬付与を繰り返し、クアッドクローラー

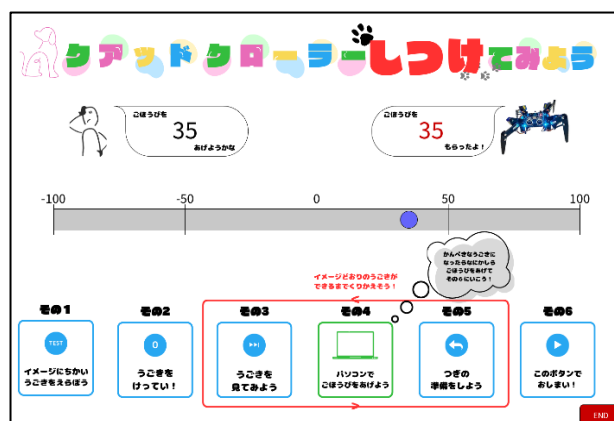


図1 報酬付与時のインターフェース

をしつける。

まず初めにしつけたい動作を考える。次に、リモコンの TEST ボタンを繰り返し押して、ランダムな動作を生成し、しつけたい動作に近いものが生成されたら 0 ボタンを押して動作を決定する。▶▶のボタンを押して強化学習により動作を生成し、生成された動作に対して PC で報酬を付与する。報酬付与に用いたインターフェースを図1に示す。報酬付与後には◀のボタンを押し、次の強化学習による動作生成を準備する。以降、イメージ通りの動作ができるまで、強化学習による動作の生成から報酬付与、次の準備までを繰り返す。最後に、適当な報酬を与えてから▶のボタンを押し、操作を終了する。

## 3. 強化学習による動作生成

本研究では、強化学習アルゴリズムとして Q 学習を採用する。Q 学習はエージェントの行動に対して報酬を与えることで目的に応じた動きを生成することができる強化学習アルゴリズムである。現在の状態で取り得る各行動に対して、当該行動

を選択した場合に得られることが期待される累積報酬を  $Q$  値として算出し、次の行動を選択する際に用いる。時刻  $t$  の状態  $s_t$  の行動  $a_t$  の  $Q$  値  $Q(s_t, a_t)$  は、学習率  $\alpha$ 、割引率  $\gamma$ 、状態  $s_{t+1}$  で得られる報酬  $r_{t+1}$  を用いて式(1)により更新される。

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha\{r_{t+1} - Q(s_t, a_t) + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1})\} \quad (1)$$

本研究では  $\alpha$  を 0.7,  $\gamma$  を 0.9,  $r_{t+1}$  の取り得る値を  $-100 \sim 100$  とした。状態  $s_t$  は 4 つの脚の付け根と膝に付属しているサーボモータの角度で表現し、行動  $a_t$  は  $\epsilon$ -greedy 法により選択する。

#### 4. 評価実験

横浜朝鮮初級学校在学中の小学 3 年生 9 名を 3 人 1 組の班に分け、本システムを操作させた後、個人ごとにアンケートを実施した。また、本システムを体験した感想をヒアリングした。アンケートでは、「イメージ通りの動きをしつけることができたか」、「ロボット自体をペットとして認識することはできたか」、「プログラミングに対する興味が湧いたか」という 3 つの項目について 5 段階で評価させた。評価ごとの回答者数を表 1 に示す。

1 班はクアッドクローラーの左前脚と右後脚を同時に上げる動作を目指し、左前脚が大きく上がる動作にたどり着いた。2 班は両後脚を伸ばした「逆おすわり」を目指し、左前脚および左後脚が伸びる結果となった。3 班は「逆立ちを覚えさせたい」という意見から、2 班と同じく後両脚を伸ばす動作を目指し、左後脚が浮き、右後脚は伸びる動作が生成でき、イメージに限りなく近い動作が生成できた。

ヒアリングでは、「ロボットを育てるのがこんなに難しいと思っていなかった」、「ごほうびを  $-100$  与えたことを後悔している」、「またやってみよう!」、「私もそのロボットを作れるようになりたい」等のコメントが得られた。

#### 5. 考察

提案手法を組み込むことで、クアッドクローラ

表 1 評価ごとの回答者数

| 質問<br>回答         | しつけ<br>やすさ | ペットと<br>しての認識 | 興味の<br>向上 |
|------------------|------------|---------------|-----------|
| はい               | 3 名        | 3 名           | 7 名       |
| どちらかとい<br>えばはい   | 5 名        | 2 名           | 2 名       |
| どちらでも<br>ない      | 1 名        | 3 名           | 0 名       |
| どちらかとい<br>えはいいいえ | 0 名        | 0 名           | 0 名       |
| いいえ              | 0 名        | 1 名           | 0 名       |

ーをペットとして班でしつけ、プログラミングに対する興味関心を向上させることができたと考えられる。一方で、ペットとしての認識がしつけやすさおよび興味の向上に比べて劣る理由は、実験中に「音を出しながら踊れるのか」、「歩く動作は学習できないのか」という意見が出ていたことから、小学生が求めている動物のような柔軟な動きを生成することができなかったからであると考えられる。また、しつけやすさの点では、最初にイメージした動作に近い動作は生成されても、理想に限りなく近い動作が生成されたのはひとつの班だけであったため、「どちらかといえばはい」という回答が多くなったといえる。

今後の課題は、プログラミングへの興味関心を高めるだけでなく、「しつけやすさ」や「ペットとしての認識」を向上させるために、強化学習を活用した動作生成の仕組みを改良することである。ユーザが入力した報酬が高いほど微調整程度の動きが生成され、低いほど大胆に動きが変わることが求められる。適切な学習率や割引率を調整することに加え、 $Q$  学習以外の手法による動作生成も試す必要がある。

#### 参考文献

- [1] 星尚徹, “クアッドクローラーを用いたアルゴリズムに対する児童の学習意欲の向上”, 東京都市大学メディア情報学部情報システム学科卒業論文, 2023.