

1. 概要

今回のゲームでは考えるボードのサイズが最大で 256x256 ということで純粋なアルゴリズムで処理するのは難易度が高いと考えた。そこで、深層強化学習を用いて現在の状態から使用するべき抜き型と座標、方向の三つの要素を推測する。

2. アルゴリズム

SAC(Soft Actor Critic)という深層強化学習アルゴリズムをベースとして開発を進めた。各ネットワークにはCNN(畳み込みニューラルネットワーク)を使用することで各ピース間のつながりを考慮しながら、与えられたボードと目標ごとに学習を行う。事前に適当な型と目標で学習を行い、初期の重みを学習に有利なように調整する。

3. 工夫点

3.1 高速化

環境(抜き型使用後のボードや報酬を出力する class)をpythonでなくC++で実装しそれをpython上で使用することで学習時間の短縮を図る。

3.2 GUI 化

型を数値ごとに色分けし、目視で目標値との差を確認できるよう GUI を作成。

4. 開発環境

Python, pytorch, c++, OpenSiv3D, Windows11