COMP20050 Assignment 5

echo

May 1, 2022

Abstract

Blokus Duo sprint 5 report. This document describes:

- a description of our bot algorithm
- a link to our Trello board showing the tasks and their allocation to team members.

Deep reinforcement learning agent

This algorithm uses a variation of the Q-learning algorithm to train a dense neural network. It uses the ai.djl library to create and train the model. The algorithm used to train the model involves passing the board as an input array into the model. In this array 1s are used to represent spaces occupied by the bot, -1s represent spaces occupied by the opponent, and 0s represent empty spaces. This observation array is passed into the model and an array of rewards for each possible move is returned. The bot sorts the array and returns the maximum valued valid move when makeMove is called. Moves are represented as one-hot encoded arrays where the index can be converted into a unique move. This makes it convenient to get multiple moves from a single model inference. To train the bot an environment that simulates the game was implemented. It runs a game and records the state, action, and score of each game step. At the end of the game the final result is propagated back through the steps to calculate the reward at each step. The actual reward for the move played is used as the target value to train the neural network with back propagation. This submission does not include a trained model due to the large size of the files.

Minimax algorithm

We also tried to implement a minimax algorithm to play the game but were unable to complete the makeMove method.

Trello Board

Our trello board can be accessed with this link: `https://trello.com/b/rTxZx2is/blokus-duo`