A Study in the Effects of Communication between AI

Yu Hirayama Kennya Kawazoe
Ayano Michiwaki Kagura Miyamoto

Abstract

We studied whether AI acting in a pair that can communicate with each other has any specific effects. We found a model which shows that there somehow was an effect caused by AI acting in a pair. However, we have not confirmed whether the effect comes from communication or not.

Introduction

When an AI solves any problem, it is normal for it to calculate with a single “brain” (which means single AI). However, we thought of the need for “cooperative AI” that can act in human society. Therefore, we aimed to design an AI that cooperates with other AI by communicating.

Method

1. Game Workflow

First, we create a pair of AI. In a 2D field, the pair will carry food to their nest in the corner. Each piece of food is placed randomly in the field, and there could be more (or less) pieces than the amount the AI can hold in each spot. The mission for them is to bring back the determined amount of food to their nest.

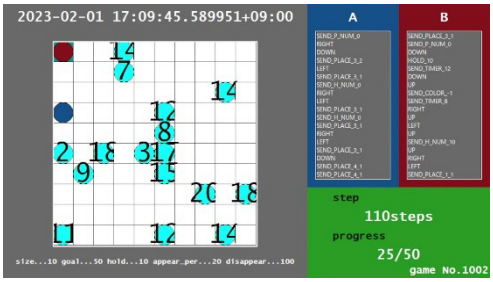
2. Reinforcement Learning

In this game, each AI is given -1 reward (which means penalty) by default each step. When they bring food to their nest, regardless of the amount, they are given +1 reward. Also, When the amount of food brought meets the quota, the AI which did not bring the last piece also gets +1 reward.

Each AI gets 15 parameters of state[1] to decide the in action. Using Double DQN, the AI estimates the value of each action in the current state. There are 10 types of action[2] that AI can are actions to move 1 square, while the rest of the actions are to send pieces of information to the other AI. Since the AI can choose either type of action at 1 step, it is important for it to choose the optimal way.

1.states	2.actions
PROGRESS	UP
MY_PLACE_ROW	DOWN
MY_PLACE_COL	LEFT
MY_PLACE_FOOD_NUM	RIGHT
MY_HOLD_FOOD_NUM	SEND_PLACE
MY_FOOD_TIMER	SEND_PLACE_FOOD_NUM
MY_FOOD_TASTE	SEND_HOLD_FOOD_NUM
MY_FOOD_COLOR	SEND_FOOD_TIMER
BUD_PLACE_ROW	SEND_FOOD_TASTE
BUD_PLACE_COL	SEND_FOOD_COLOR
BUD_PLACE_FOOD_NUM	
BUD_HOLD_FOOD_NUM	
BUD_FOOD_TIMER	
BUD_FOOD_TASTE	
BUD_FOOD_COLOR	

※Data in colored boxes are dummy data.



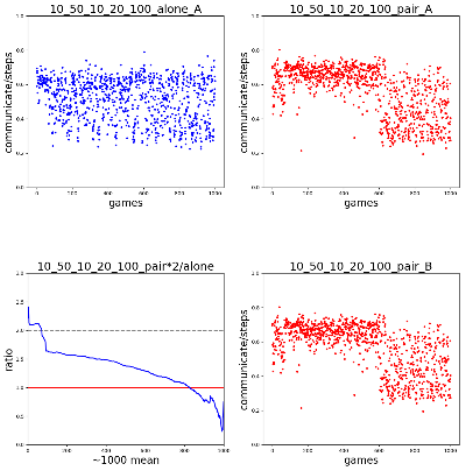
※Image of the game field.

Analysis

To analyze the AI's learning, we used this formula.

$$E = \frac{(Average\ steps\ when\ played\ in\ a\ pair) \times 2}{(Average\ steps\ when\ played\ alone)}$$

E comes from "effect", and when $E<1$ is true, it means that there somehow was an effect that made the AI bring food more efficiently. This time, we found a model that turned out to be lower than 1. In this case, according to the scatter plot, we guessed that the AI which acted in a pair tended to choose moving and communicating with less randomness. However, this was not true in all models. In some model E was Higher than 1.



※Versions of the program

Python 3.10.5

keras 2.9.0

numpy 1.22.4

pandas 1.4.2

tensorflow 2.9.1

pyinstaller 5.7.0

(Using pipenv==2022.12.19 to create exe file)

※Windows 11 Home 64bit

Outlook

In this study, we still have not found a general tendency. Since there are many parameters to adjust, we are thinking of producing more models. Also, we will collect more detailed data to analyze the relationship between each action.

References

[1] 高校数学からはじめるディープラーニング 初歩からわかる人工知能が働くしくみ(金丸 隆志 , 講 談 社 ,2020.4.16 発 行 ,ISBN978-4-06-519403-4)

[2] 現場で使える！Python深層強化学習入門 強化学習と深層学習による探索と制御(伊藤 多一/今津 義充/須藤 広大/仁ノ平 将人/川崎 悠介/酒井 祐 企 / 魏 崇 哲 , 翔 泳 社 ,2019.08.07 発売,ISBN978-4-7981-5992-8)