Learning to Escape From a Dungeon Maze: A Game AI Player via Deep Reinforcement Learning

s1260086 IKEDA Masaki, Supervisor: LI Xiang

1. Summary of the Research

- This research is conducted on Unity using ML-Agents. ML-Agents is a Deep Reinforcement Learning (DRL) framework provided by Unity.
- AI agent aims to escape from dungeon like a maze. Additionally, there are some enemies in that dungeon.
- The goal of this research is to have agent learn how to avoid enemies and escape from dungeon with sensors.
- There are previous researches such as agent learns to solve maze without enemies [1] or to defeat enemies [2]. In this research, agent should solve the escaping problem with enemies without defeating enemies. Agent can only avoid to be caught by enemies. That is a difficult point.
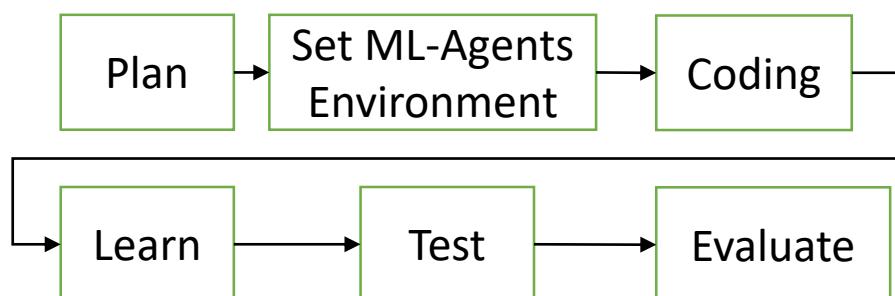- This research is using dungeon provided by Unity Assets that named John Lemon Haunted Jaunt.

Fig 1. Flow of the research

2. Methodology

- The agent is trained by DRL. PPO (Proximal Policy Optimization) algorithm is used for this research. It can solve complex formula than TRPO. Algorithm is able to be simply implemented in the environment.
- ML-Agents is framework for constructing learning environment of machine learning on Unity. By using it, this study can program by *C#* and consider about reward system and hyperparameter.

Fig 2. Main learning environment
(Main dungeon. The escape route is with green color.)

3. Experiment and Results

3.1 Knowledge Prepared

- Deep Learning from Scratch
- Introduction and Programming of Deep Reinforcement Learning (DRL) Algorithms
- Unity ML-Agents Environment

3.2 Project Progress

- Successful pre-training on stage-by-stage maps (divide as functional pre-training area.)
- AI agent can escape from dungeon without enemies.
- AI agent can escape from dungeon with enemies.
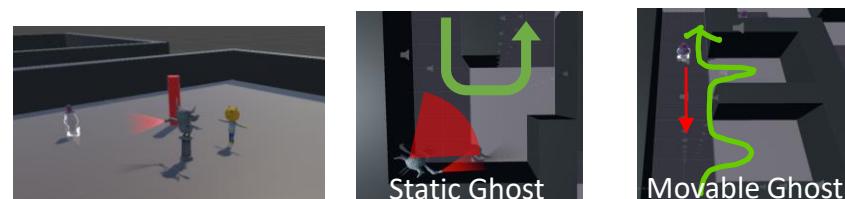

Static Ghost   Movable Ghost

Fig 3. Pre-training area and learning stages

- Main problems and the solutions:
  - First, the agent tends to escape quickly, so the agent cannot often wait for passing through enemies. This problem can be improved by implementing "DamageZone" where the agent obtains negative reward.
  - Second, the agent is sometimes caught by enemies after turning at corner. The "Risk Zone" or "Sound Zone" is considered.

3.3 Evaluation criteria

- In this research, the performance is evaluated by the success ratio of escaping as following formula.

$$success\ ratio = \frac{Escape\ count}{Trial\ count}$$

  - Escape count is the number of times that agent was able to escape from main dungeon.
  - Trial count is the number of times that agent tried main dungeon.

4. References

[1] Yiou Wang and Yujie Wang. "Moving Architecture, Animated Maze: The Intertextuality Between Player and Environment Agents." *Creativity and Cognition,* 2021.

[2] Adamsson Marcus. "Curriculum Learning for Increasing the Performance of A Reinforcement Learning Agent in A Static First-Person Shooter Game," KTH Royal Institute of Technology, 2018.