Group No:15

Ege Berkay ,İbrahim Taha Aksu

Erdem Ünal ,Onur Kulaksızoğlu

Burak Gök

Cs-464 Project Proposal

1 Problem Description

Problem we have in hand is creating a self learning agent in a small environment of the game Doom. This is a special map where the character cannot move but just turn around and shoot. Agent will learn to defend himself by killing the monsters coming to it from around. Monsters will respawn with more health after they are killed. Moreover the agent will have limited ammo of 26 to prevent spamming. The allowed actions for the agent are turning left, turning right and shooting. The game ends when the player dies or the time limit of the game is reached which is 60 seconds total.

2 Approach

We are going to implement reinforcement learning to train our agent. Reinforcement learning allows machines to follow some behaviors given feedback on their past actions. This approach has basically makes use of 2 mechanisms to decide the current action: first a system to describe the current state of the environment, second a system that decides on the action given the current state and given a defined policy.

The determined rewards for the agent: +2 points for killing a monster, -1 for dying and -10/26 points for failing to kill a monster after shooting(to prevent wasting ammo).

Let us know explain our algorithm in detail as it was addressed during our presentation. For the first part, that is to describe the environment to our agent, we will be using convolutional neural networks which is a class of neural networks that has successfully applied to analyzing visual imagery. The second part which takes state and policy definition and in return decides on the action is called Q-learning. Thus in sum we will have our environment feed us the frames from the game which will go through the CNN tree and become state definitions. This state information will be fed to Q learning with the policy we defined and Q learning technique will decide on the new action to be taken.

3 Data

OpenAI gym environment will be used for obtaining training data. The data will extracted from randomized games played with DoomDefendcenter-v0 environment.

4 Evaluation

We aim 10 points on average out of 100 consecutive games as our success limit. OpenAI gym has some other solutions on the problem with whom we can compare our algorithm too. If we manage to finish this problem early we can implement same approach to other harder/unsolved problems again in a doom environment.

5 Work Statement

Experiments with OpenAI environments/library and I/O: 5~ hours: Berkay, Onur

Researching algorithms, reading papers: 5~ hours: Taha, Burak, Erdem

Programming / Testing iteratively: 10~hours: Altogether. Total of 75~ man-hours