Seungchan Yang

December 10, 2018

CS395

Reinforcement learning outcome

Reinforcement learning is an area of machine learning. Learning process uses an infant agent to take an action to learn the optimal way of making decisions. For example, there is a famous computer program called AlphaGo. This program learned the optimal way to take an action for playing Go which is an abstract strategy board game for two players. At the first the program did not know a good way to play this board game, but after experiencing games with other players, AlphaGo learned the best way to play this game. For the reinforcement learning, I used the unity environment and made a small 2D top down game. To make a game using the unity, I have to learn C sharp. C sharp is one of object oriented programming language. Grammar is almost same as c++, therefore it is not hard to implement functions or algorithm.

For implementing reinforcement learning in a game, there are three important keys which are reward, action, and state. In the 2D game, there is an agent moving around in a map. This agent moves top, down, right, or left. At the first the agent randomly chooses direction, because it does not know which action is the best action. After taking an action, each state gets a reward. If an action makes to the exit, then the state gets high reward from it. If not, the state gets negative reward. The state is a dictionary which as a key value and state values. The key value can identify specific state; therefore, the program can save learning data in a correct place. The purpose of agent is to maximize the sum of rewards. At the beginning of the reinforcement learning process, the agent will move randomly, because it does not know which action is the best. However, after each state gets the sum of rewards then the agent can choose the best action. The state represents each place of the map. The size of map is eight by eight. State keys are can be 0,0 to 7,7.

The program saves data dynamically. Therefore, after the agent learns the values of each actions, it can choose the best action. The agent will choose better action as much as it moves. The reinforcement learning will get more accurate as time goes by. The algorithm that is used in this program is called Q-learning algorithm. After taking an action, the program measures rewards and update Q values. Q values are the state values that gives information to the agent for the optimal way to take an action.

Through this project, I learned how to implement reinforcement learning and how to use unity to make a game. This was one of great experience in college. I had experienced with making a roguelike game. Also, I liked to play computer games. Therefore, I chose this course. I expected that this course would give me a knowledge to make a game. However, this course gave me more than just making a game, but also it gave me experience with AI implementation. I never had experience with machine learning and the unity environment. Therefore, I thought I would have hard time for this project, but throughout this course, Dr.Harrison helped and taught me. Therefore, I could learn the basic algorithm of machine learning. I believed this learning outcome could be used in various computer science fields.