Overwatch Reinforcement Learner

Max Leach

Overview:

Overwatch is a strategic team-based shooter that rewards more on “game sense”, or the ability to act accordingly in a given environment, over mechanical skill such as aiming. There are, at the time of the creation of the agent, 30 different heroes; each hero has their own playstyle and abilities, furthermore, learning all of them would be too computationally resourceful and so the agent will only be playing one hero, Tracer. The aim is to create a reinforcement learning agent that can play in a 6v6 team-based environment and account for other teammates while play accordingly relative to its environment.

Problem

OpenAI, an AI focused company thriving to create safe general intelligence, have created 6 agents in the video game Dota 2, another strategy required game, using the model PPO (Proximal Policy Optimization) and no direct links to each other’s PPO model meaning over time, through self-improvement has been able to co-operate to achieve team cohesion; while there are indirect links between the models, it can co-operate because there is a parameter input known as “team spirit” which is a second factor/input into the model (in the PPO model, the other input is an LSTM model which predicts the most appropriate action to gain as much rewards through the input of previous timesteps). OpenAI were able to achieve such high-level performance from the bot as they trained the agent through 180 years of experience per day, compared to the human learning ability, and were able to achieve 180 years of experience per day through a supercomputer consisting of 128,000 CPU cores and 250 GPUs. Furthermore, I will not have access to that computational power, even if I had a model with the capability of playing all the hero roster at a mechanically and strategically level superior to which of a professional player, it would be unachievable. Another problem is pynput (a python library to input mouse/keyboard strokes), while it can emulate keyboard inputs, it cannot for mouse controls (possibly used as anti-bot detection, to prevents aim cheating) an alternative for pynput would be an external program running alongside the python main LSTM model written in java to bypass libraries and to directly input win32 commands; an external program may also slow down the latency and may be too slow to be able to decipher frames from recordings in real time.