Training Artificial Intelligence Agents to Play a Tower Defense Game Using Reinforcement Learning

Mitchell Harrison, Jonathan Rivera, Justin Stevens, and Dr. Chad Hogg

Spring 2024
Computer Science

Abstract

The tower defense video game genre poses interesting challenges for Artificial Intelligence (AI) training. An AI must keep track of a large amount of in game data and make numerous decisions in a brief time span. We developed a tower defense game within the Unity Game Engine and used the Proximal Policy Optimization (PPO) Reinforcement Learning algorithm to train the AI to play the game.

Bloons Tower Defense

Tower defense (TD) is a genre of video games, requiring players to place defensive structures, known as towers, to stop enemies from reaching an endpoint. One notable example of TD is Bloons Tower Defense (BTD), where players face off against colorful balloon enemies known as "bloons" and must use various monkey towers to pop them before they reach the end of the path. We recreated this game.

Reinforcement Learning

Reinforcement Learning agents repeatedly observe the game state, choose an action, and receive a reward or punishment. Over time, they learn to choose the actions that produce the greatest cumulative rewards. We used OpenAI's PPO reinforcement learning algorithm. We needed to determine the state and action representations and the reward scheme for our game.

AI Implementation

Observations: Money, Lives, Current Wave, and 16x10 map (160 tiles).
Actions: Do nothing, buy a tower (Dart Monkey or Sniper Monkey), X and Y position of where to place a tower, and tower targeting mode (first, last, or strongest).
Rewards: Winning all the waves (1), winning a wave (0.5), and incorrectly placing a tower (-0.01).

Figures

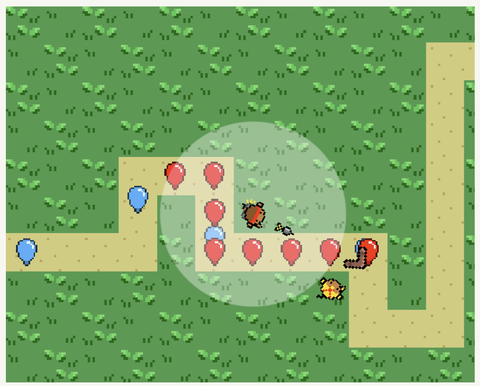

Figure 1: In-game screenshot showing bloons being popped by two monkey towers.

Figure 2: In-game screenshot showing a small subset of AI agents training.

Results

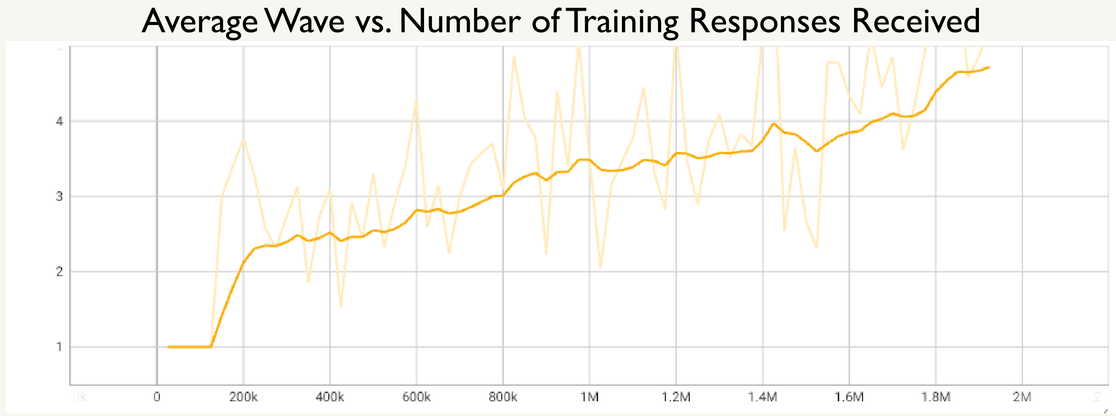

Figure 3: Results of AI getting to higher waves as it learns over 4.2 hours. Faint orange line shows the actual data while the dark orange shows the smoothed trendline.