Neural MMO: Ingredients for Massively Multiagent Artificial Open Worlds

Joseph Suarez, Phillip Isola

Project Page
jsuarez5341.github.io

Community Discord
discord.gg/BkMmFUC

Source Code
jsuarez5341.github.io

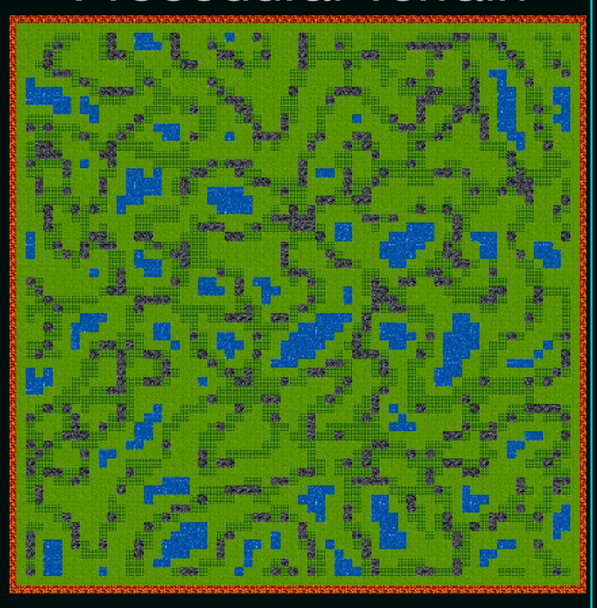
Twitter
@jsuarez5341

Abstract: Simulated games have become a staple of multiagent intelligence research. Our work considers massively multiplayer online role-playing games (MMORPGs or MMOs), a genre of games that has only recently begun to gain attention in the reinforcement learning community. MMOs capture several complexities of real-world learning that are difficult to integrate with existing methods. We present several highlights from the ongoing development of Neural MMO that are particularly relevant to enabling reinforcement learning methods in artificial open worlds.

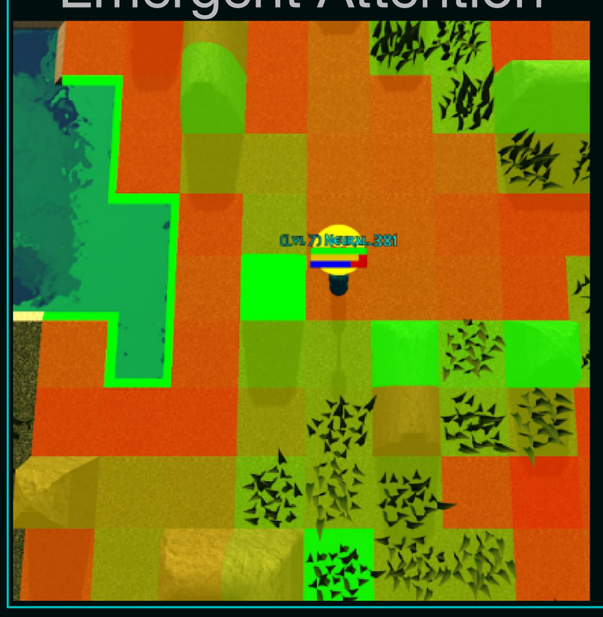
Neural MMO v1.4 Baselines

Persistent Skills

Procedural Terrain



Emergent Attention



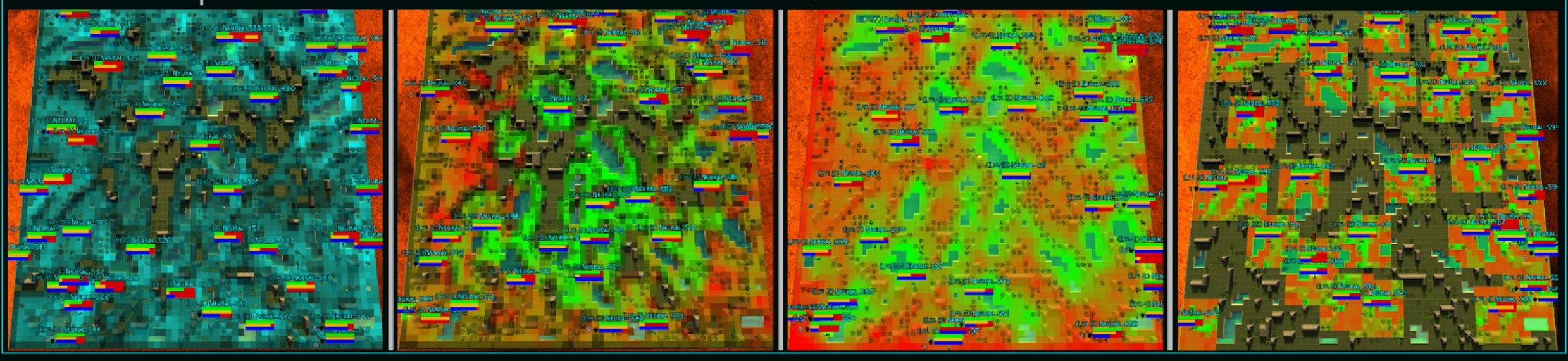
Learned Combat

Baselines: Each policy was trained overnight with four CPU cores and a single GTX 1080 Ti GPU at 10% utilization using RLlib's PPO implementation with default hyperparameters. Model sizes are 620k, 175k, and 140k (attentional, convolutional, recurrent).

The distribution of lifetimes is long-tailed because new agents must evade high-level adversaries. Early success leads to robust foraging and combat over thousands of timesteps. We have seen individual agents live for over an hour and a half in real time (10k+ timesteps)

Overview: Neural MMO agents inhabit persistent, procedurally generated worlds with randomized food, water, and obstacle distributions. Individual agents progress various foraging and combat related skill levels throughout their lifetimes. Over the course of training, agents develop better strategies and learn to reach higher skill levels. In-game overlays allow us to visualize the various meaningful quantities that agents have learned about their environment. For example, agents have learned that food, water, and nearby obstacles merit their attention while other tiles do not.

Counts Explorations | Local Value Function | Global Value Function | Local Attention



Overlays: Even among multiagent tasks where performance is always highly dependent upon the actions of other agents, open worlds are uniquely difficult settings for measuring policy quality. Neural MMO and the game genre that inspired it are not zero-sum, are neither fully cooperative nor fully competitive, and defy most attempts to impose ELO metrics often used in more strictly specified multiagent tasks like DoTA and Starcraft 2. We propose in-game overlays to mitigate difficulties in interpreting learned policies. Counts exploration shades tiles based on the number of visitations. Baseline policies explore all but a few small and resource-poor areas of the map. Shading tiles based on value functions shows that agents have learned to value terrain with both resources rather than food or water individually. The attention overlay shows which tiles most influence agent actions and shows that agents have learned to focus on food, water, and nearby obstacles.