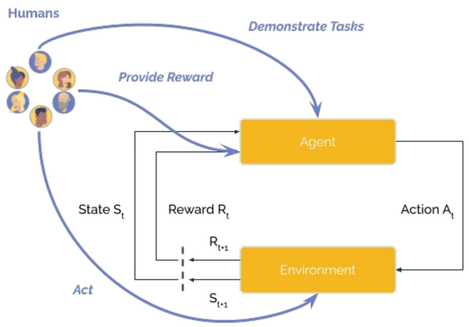
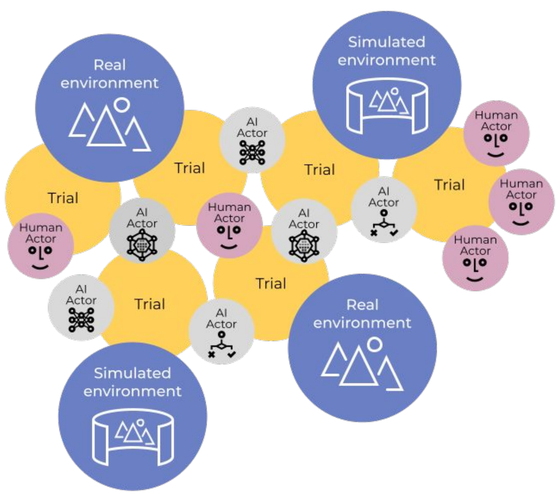
Cogment: Open Source Framework For Distributed Multi-actor Human-AI Collaborative Environment

Sai Krishna Gottipati, Clodéric Mars, Gregory Szriftgiser,
Sagar Kurandwad, Vincent Robert and Francois Chabot

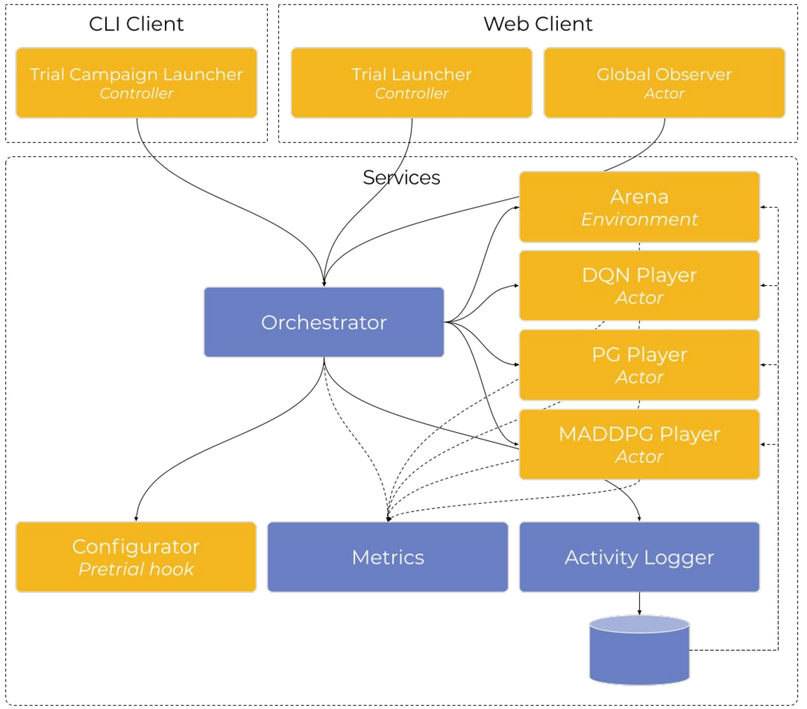
Motivation:
To develop social intelligent behaviours, humans need to react with agents in multiple ways: acting in the environment on par with agents, providing rewards for the agents, demonstrating tasks (e.g, imitation learning), designing tasks (e.g, curriculum learning) etc..



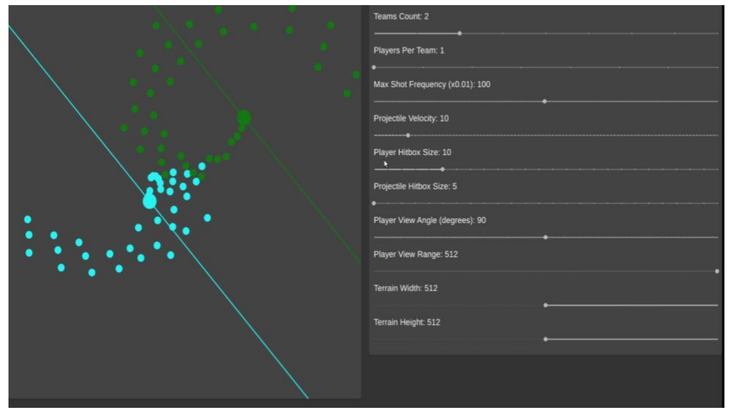
Our proposed framework Cogment (based on micro-service architecture) orchestrates the running of trials involving AI and human actors in simulated or "real-world" environments. Each Trial involves one or multiple actors and runs in one environment. Multiple trials can run concurrently and can share actors or environments.



To view the full documentation and other resources, visit cogment.ai



Death match was developed as a testbed and showcase for multi agent reinforcement learning (MARL) using Cogment. The game is a competitive and cooperative paintball-like shooter, where teams of agents compete against each other in an arena. Agents shoot paintballs at their opponents from other teams. An agent is eliminated if it is hit by an opponent; there is no friendly fire. Last team standing wins the game; each Cogment's trial consists of a game. It is now adapted as a social intelligence benchmark in a lifelong learning setup similar to that of Hanabi. Death match includes a lightweight web based front-end where users can configure and watch a trial. Figure below features a duel between two agents trained using self-play. (The larger circles are the players, the smaller circles are the paintballs)



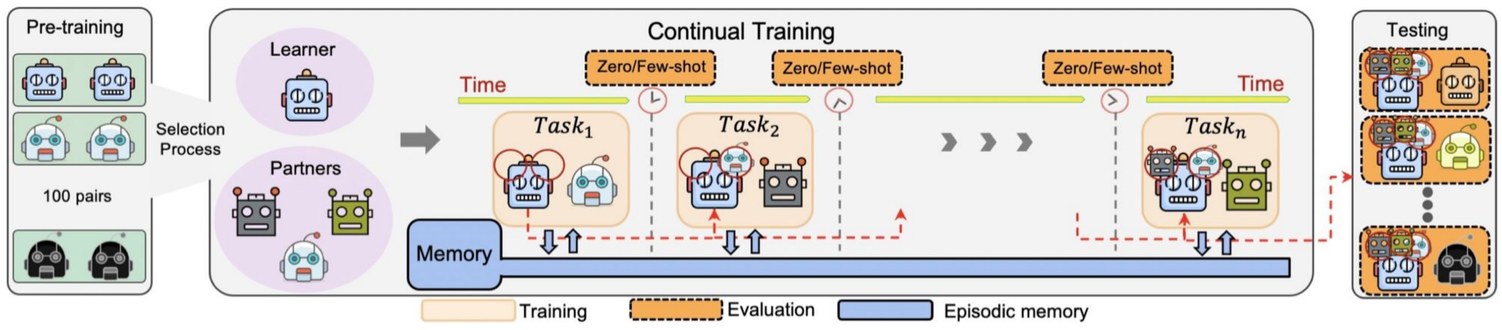


Image source. Similar to Nekoei et al, our social intelligence benchmarking task on Death Match involves three phases. In the first phase, different agents (with varying algorithms, architectures and seeds) were trained using self-play. Their performance in this phase was judged based solely on the rewards achieved at the end of their training. In phase-2, agents were sampled from this trained pool of agents and paired against each other. The agents are expected to develop social intelligence behaviors in this phase. They are scored based on several different metrics including total survival time of each of the agents, number of paintballs fired by each agent, collaborating with other agents to fight a common enemy etc.. In phase-3, each agent is made to play against agents (and humans) that it hasn't played before