Final Project Proposal

Jon Reasoner, Tyler Kim

Motivation/Background: Teaching a robot to understand human behavior and predict future action is a critical task, but an exceedingly challenging one. Reinforcement learning is an approach that can train robot control policies, optimizing decision-making based on their state and the state of human actors in their vicinity. One such example of this is moving a robot through a crowd. A robot should be able to view the behavior of the human actors and act in a manner that ensures that a collision does not occur. Furthermore, scenarios exist where you not only need to avoid obstacles but also avoid detection. In a scenario where this is true may be in supply delivery to humans in an environment where hostile actors are trying to prevent that supply delivery such as authoritarian regimes or battlefields.

Problem Definition: This project aims to further expand on the previous example of obstacle avoidance, where the obstacles are humans, to not just avoid the humans but avoid being detected by the humans. This is like the childhood game, "cops versus robbers," where the cops (humans in this scenario) are trying to find the robber (robot) while the robber attempts to go from point A to point B. The robot will start with a full map of the environment and the pose information of the humans and be tasked with creating a trajectory that enables both the avoidance of the humans and their gaze. The humans will be dynamic and moving in a semi-random pattern. The scenario will gradually be scaled up in complexity from a simple environment to an environment with obstacles, and we will attempt to implement it on hardware in the Autonomous Mobile Robotics (AMR) Lab.

Direction of Solution: Initial literature reviews point to Proximal Policy Optimization being a promising avenue for solving this solution.

Robot/Simulator for Use: We will initially attempt to use Gazebo for 2D implementation. After successful implementation, the complexity of the environment will be increased. Once successful we will try to implement it in the AMR lab using ROSBOTs or Turtlebots. Finally, we will attempt to implement in AirSim for a quadcopter, if time allows. Depending on the success of the UAS simulation, we can attempt to implement it on Crazyflie quadcopters in the AMR lab as well. During the physical implementation of the control algorithm that is trained in simulation, the humans will be simulated by Jackel robots, and the "robber" will be a Turtlebot or ROSBOT. Position data will be obtained from the lab Vicon local positioning system and sent directly to the "robber."

Target Resources:

Almazrouei K, Kamel I, Rabie T. Dynamic Obstacle Avoidance and Path Planning through Reinforcement Learning. *Applied Sciences*. 2023; 13(14):8174. https://doi.org/10.3390/app13148174

Choi J, Lee G, Lee C. Reinforcement learning-based dynamic obstacle avoidance and integration of path planning. Intell Serv Robot. 2021;14(5):663-677. doi: 10.1007/s11370-021-00387-2. Epub 2021 Oct 6. PMID: 34642589; PMCID: PMC8493784.

TorchRL PPO Tutorial : https://pytorch.org/tutorials/intermediate/reinforcement_ppo.html