

Thesis

Defense

Computer Science Master's Program

“Adversarial Deep Reinforcement Learning for Tank Duel Simulation Using Lidar-Based Observations”

By Braedan Kennedy

Abstract:

Previous research has demonstrated that reinforcement learning agents can learn to steer differential-drive robots around obstacles using 2D lidar scans as observations. However, these studies typically treat all range returns as undifferentiated obstacles—objects to avoid—without distinguishing between different object types. This thesis builds upon previous research by introducing an adversarial task in which an agent must interpret raw range readings to both avoid static obstacles and identify, pursue, and engage a hostile target.

To investigate this problem, this thesis introduces TankGame, a novel, lightweight 2D tank duel simulator. Each agent receives a 360° lidar scan, controls its motion via tread velocities, and fires slow-moving projectiles within arenas containing either procedurally generated or handcrafted obstacle layouts. Obstacles are modeled as circles of varying radii, while tanks and projectiles have rectangular profiles, requiring agents to infer object types purely from geometric cues in the lidar data. Implemented in C++ with a standardized Python API, TankGame achieves approximately 4,000 simulation steps per second on a single CPU thread—offering a high-fidelity yet computationally efficient alternative to grid-world abstractions and complex game engines.

Alongside the TankGame environment, this thesis presents a three-stage training methodology. In the first stage, a Proximal Policy Optimization (PPO) agent with a Circular Convolutional Neural Network (CCNN) encoder is trained against a static opponent, allowing it to jointly learn lidar features and a baseline policy. In the second stage, new actor and critic heads, each incorporating a Long Short-Term Memory (LSTM) layer, are trained atop the frozen CCNN encoder, enabling the agent to exploit temporal dependencies within the lidar features. Finally, in the third stage, the LSTM-PPO agent undergoes adversarial fine-tuning through self-play to improve its robustness against adaptive opponents.

Agents are trained on procedurally generated maps and evaluated on a suite of three evaluation map sets. A batch of 400 procedurally generated maps assess overall performance, while two collections of handcrafted maps test obstacle navigation and the ability to locate and eliminate static or moving targets. Experimental results demonstrate the LSTM-PPO agent defeats the PPO baseline in 64% of head-to-head duels (baseline wins 20%) and can locate and eliminate static targets in 97% (baseline 88%) of the procedurally generated maps, with additional gains achieved through adversarial fine-tuning.

The key contributions of this work are the open-source TankGame environment, a reproducible training methodology, a suite of evaluation maps, and high-performing baseline agents to support future research in adversarial reinforcement learning.

Date: Tuesday, June 3rd, 2025

Time: 9:00 AM – 11:00 AM

Location: 14-232b

Committee: Dr. Seng, Dr. Ventura, and Dr. Canaan