PROJECT 2 - 517

Summary of RL Training Tasks:

The Initial Bad Hockey Player:

The agent was set up using a PPO configuration. It began performing random actions, doing poorly, but this helped to lay the template for further changes and visualization. The focus was on letting the agent explore and become familiar with basic movement and action feedback.

SEE: SLIDE 3 in “G3 - GIFS - Project 2”

Fixed Goal:

The agent was trained with a fixed starting position and a consistent goal location. The training taught the agent to move the puck toward the goal efficiently, optimizing rewards for reaching the target. After more training, the agent started improving and began scoring more consistently.

SEE: SLIDES 3-6 & 8 in “G3 - GIFS - Project 2”

Random Start Goal:

The random starting positions for both the puck and the agent were implemented for the agent to learn to adapt to varied goal position scenarios. The training process had focused on handling different setups while still scoring on the goal. The agent developed more flexible strategies and could perform reliably under the varying conditions.

Reflection on Failures:

1) Agent Ignoring the Puck:

At first, the agent completely ignored the puck and just wandered around, the reward function didn’t push the agent to actually engage with the puck. Fixed it by adding rewards for getting closer to the puck.

SEE: SLIDE 7, gif 1 in “G3 - GIFS - Project 2”

2) Agent Getting Close to the Puck and Backing Away:

Sometimes the agent would get close to the puck but then back away or just circle around it, due to the reward function not penalizing wasted movement. By tweaking the rewards and refining the action mapping, the agent sticks with the puck and moves toward the goal more effectively.

SEE: SLIDE 8, gif 2 in “G3 - GIFS - Project 2”

3) Agent Getting Close to Scoring (Fine-Tuning):

As the agent got better, it started getting the puck near the goal but didn’t quite score or hit the goal-post. Added extra rewards for reducing the puck’s distance to the goal and fine-tuned the training with more exploration.

SEE: SLIDE 8, gif 3 in “G3 - GIFS - Project 2”

Reflection on Key Changes:

One of the biggest breakthroughs came from improving the reward function and refining the observation space. By rewarding the agent for moving closer to the puck and the goal, there was a noticeable improvement in its behavior. Updating the observation space to include details like puck velocity helped the agent make better decisions in the environment.

If we could do it differently, we’d prioritize creating a stronger reward system and focus on optimizing hyperparameters earlier in the testing process. Exploring more advanced techniques for exploration and training stability might have also saved time and avoided some of the struggles we faced throughout.

SEE: SLIDE 9, in “G3 - GIFS - Project 2”