

TraffiQ

Smart Traffic Signal Optimization

Project Overview

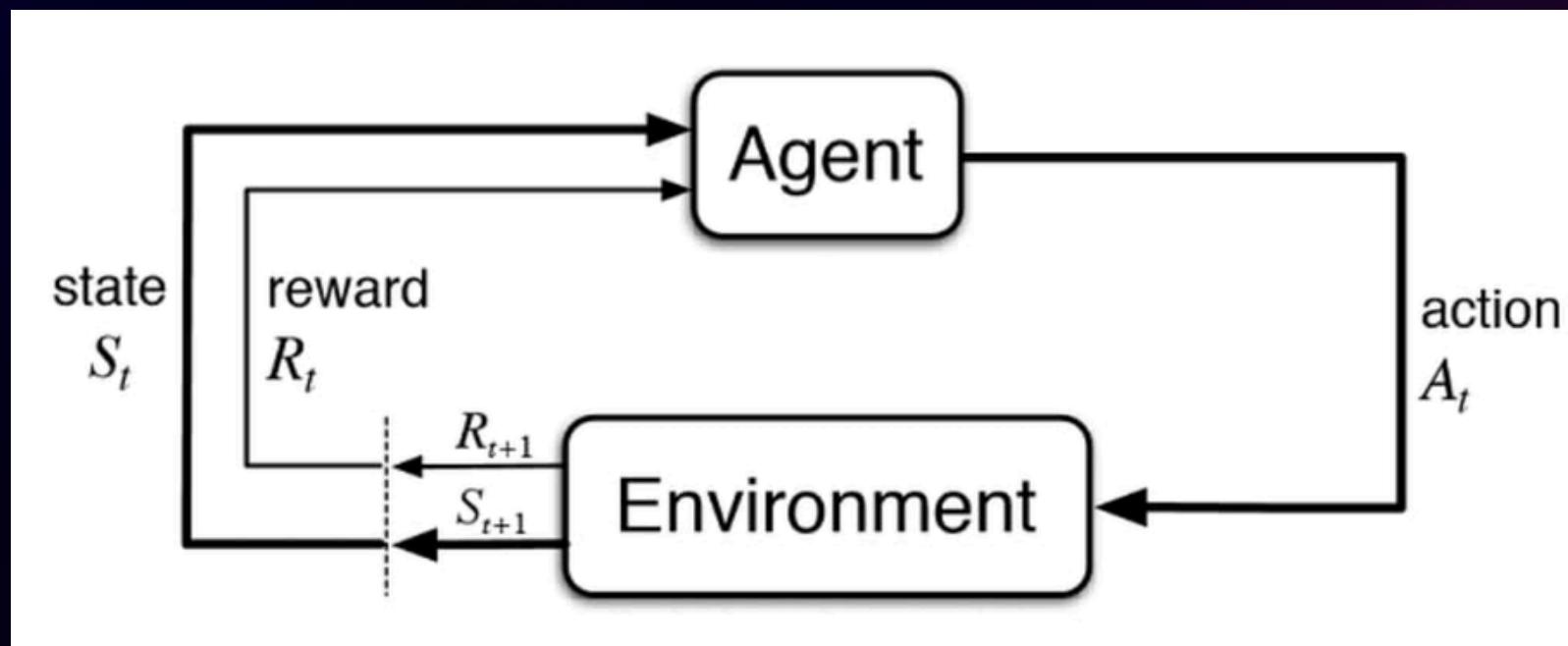
- TraffIQ aims to develop a **dynamic traffic light control system** using Reinforcement Learning integrated with SUMO.
- The RL agent learns optimal signal policies based on real-time traffic states to minimize average wait time and improve throughput.
- Starting with a single intersection and scaling to complex networks.

CURRENT PROGRESS

RESOURCES:

Reinforcement Learning Course by David Silver

Deep Learning and Neural Networks Introduction Course
1 by Andrew NG



01 REINFORCEMENT LEARNING

An agent interacts with an environment to learn a policy that maximizes long-term reward via trial-and-error.

02 MDPS & VALUE FUNCTIONS

RL is formalized as a Markov Decision Process (MDP). Agents aim to learn the optimal policy (π^*) using value functions (V , Q) governed by the Bellman Optimality Equation.

MODEL-FREE PREDICTION (TD LEARNING)

- Learns value functions without knowledge of the environment's transition model.
- Uses bootstrapped updates (e.g., TD(0)) to estimate $V(s)$ from sampled transitions.
- Balances bias and variance better than Monte Carlo methods.
- Enables online, incremental learning during agent-environment interaction

MODEL-FREE CONTROL (SARSA & Q-LEARNING)

- Learns optimal policies without a model by directly interacting with the environment.
- SARSA (on-policy): updates $Q(s,a)$ using the action actually taken.
- Q-Learning (off-policy): updates $(Q(s,a))$ using the maximum estimated value of the next state.
- Suitable for large-scale problems where modeling transitions is impractical.

TAXI: The Multi Objective Open AI Gym Env

Action Space

- 0: Move south (down)
- 1: Move north (up)
- 2: Move east (right)
- 3: Move west (left)
- 4: Pickup passenger
- 5: Drop off passenger

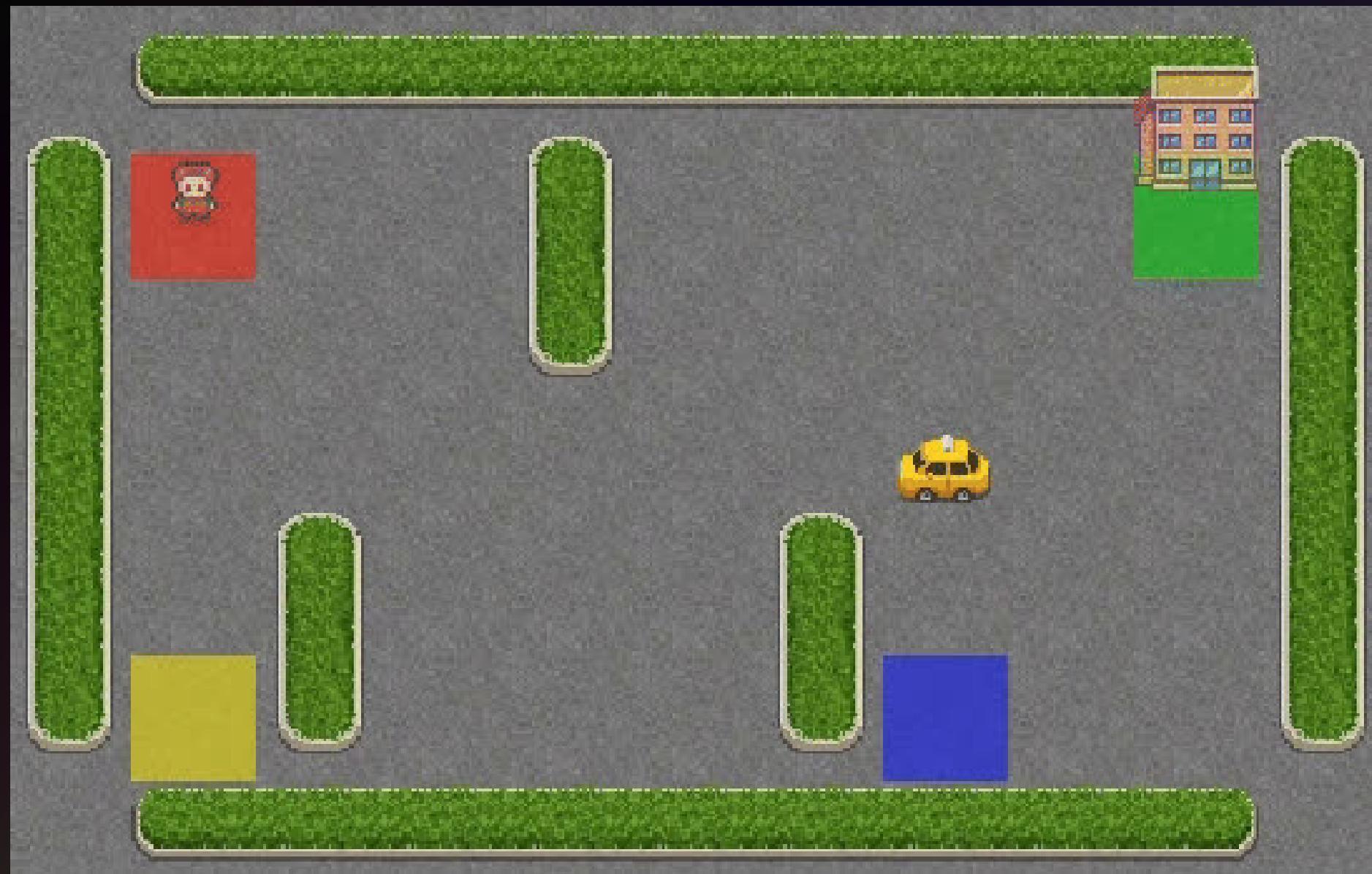
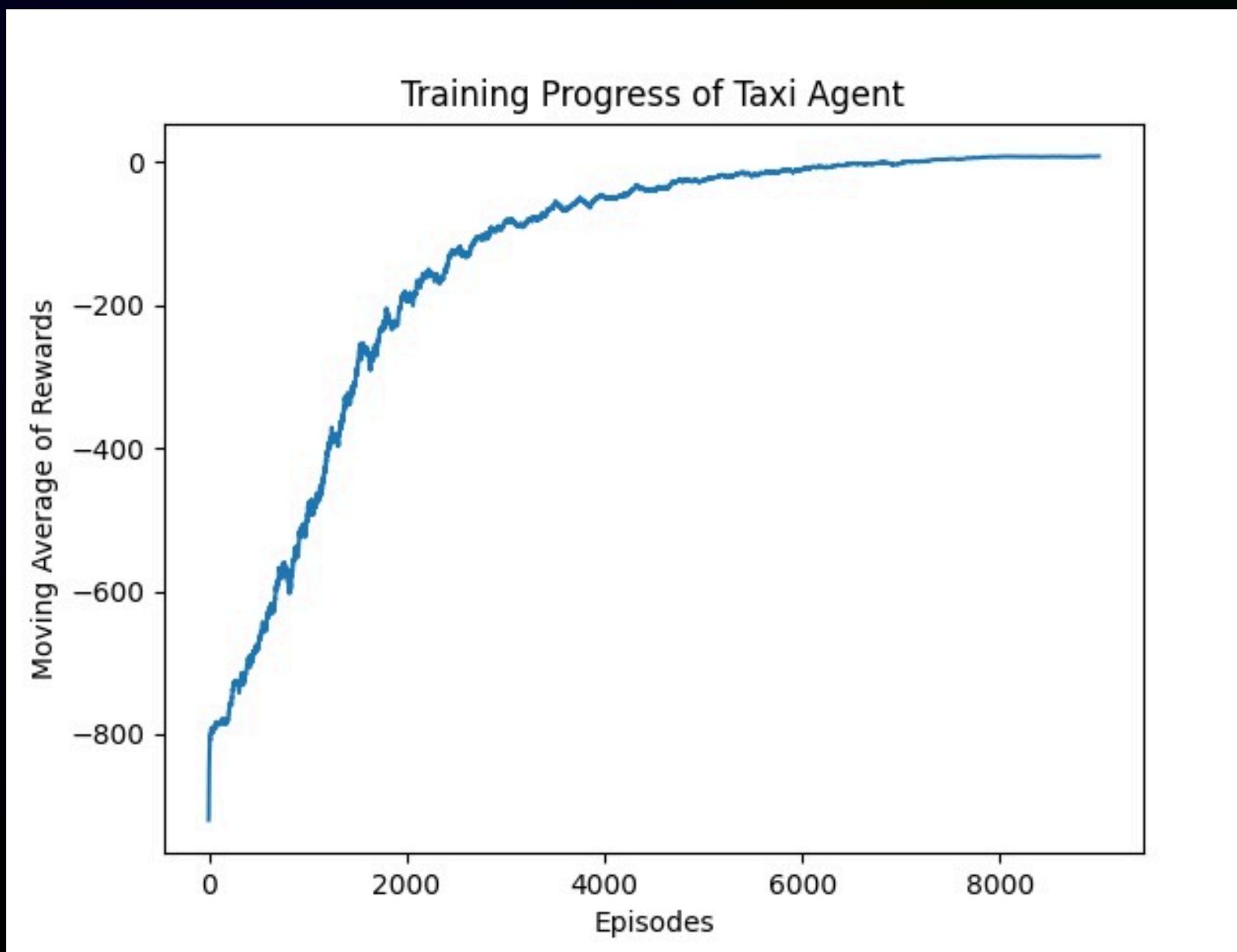
State Space

250 Discrete States
-5X5 Grid for Taxi Loc
-5 options for passenger
-4 options for apartment

Rewards

- -1 per step
- +20 delivering passenger.
- -10 “pickup” and “drop-off” actions illegally.

Performance Evaluation



Frozen Lake (Slippery) OpenAI Gym Env

State Space

4X4 Grid
-Goal at 16th cell
-4 Holes
-16 Discrete States

Reward Structure

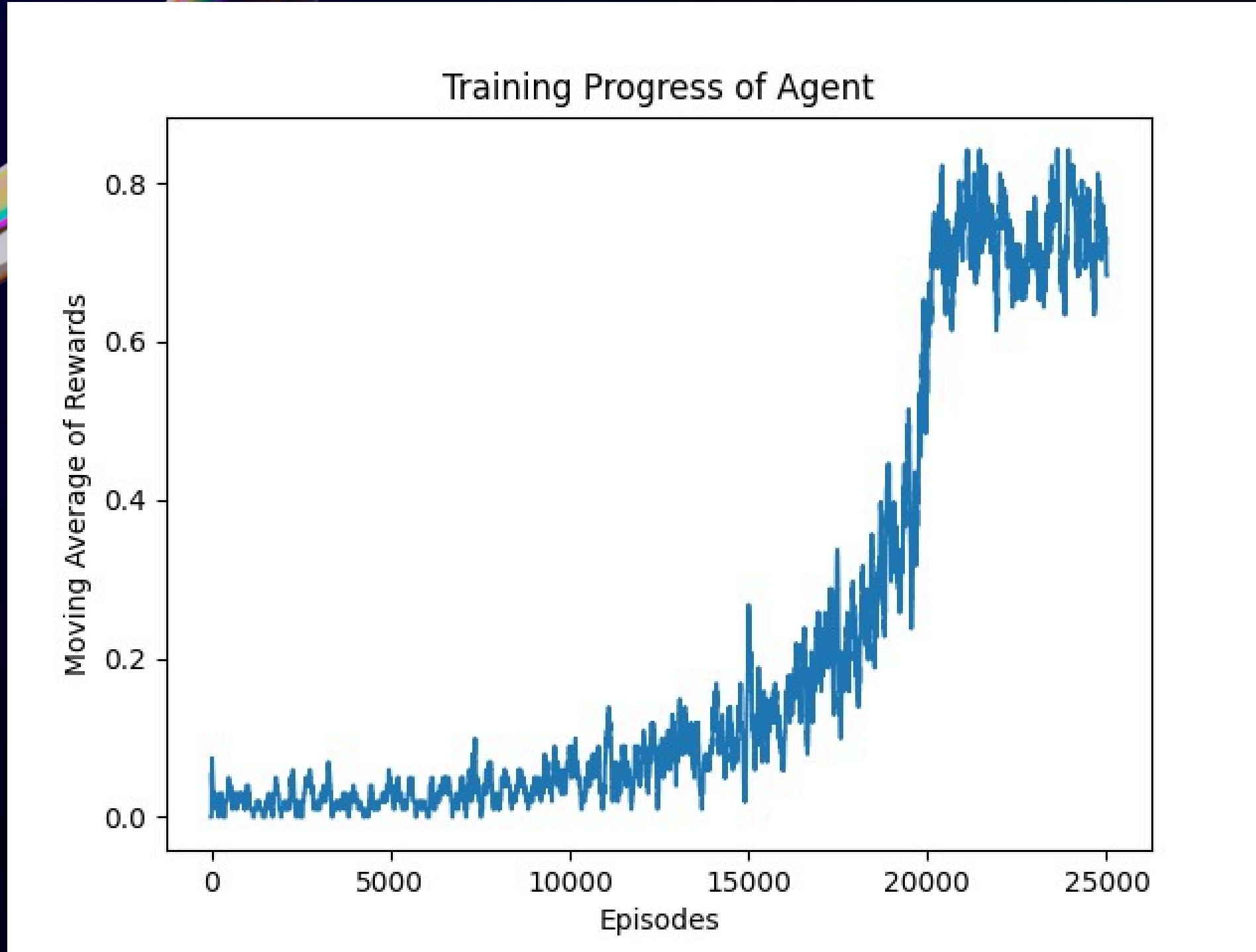
- Reach goal: +1
- Reach hole: 0
- Reach frozen: 0

Action Space

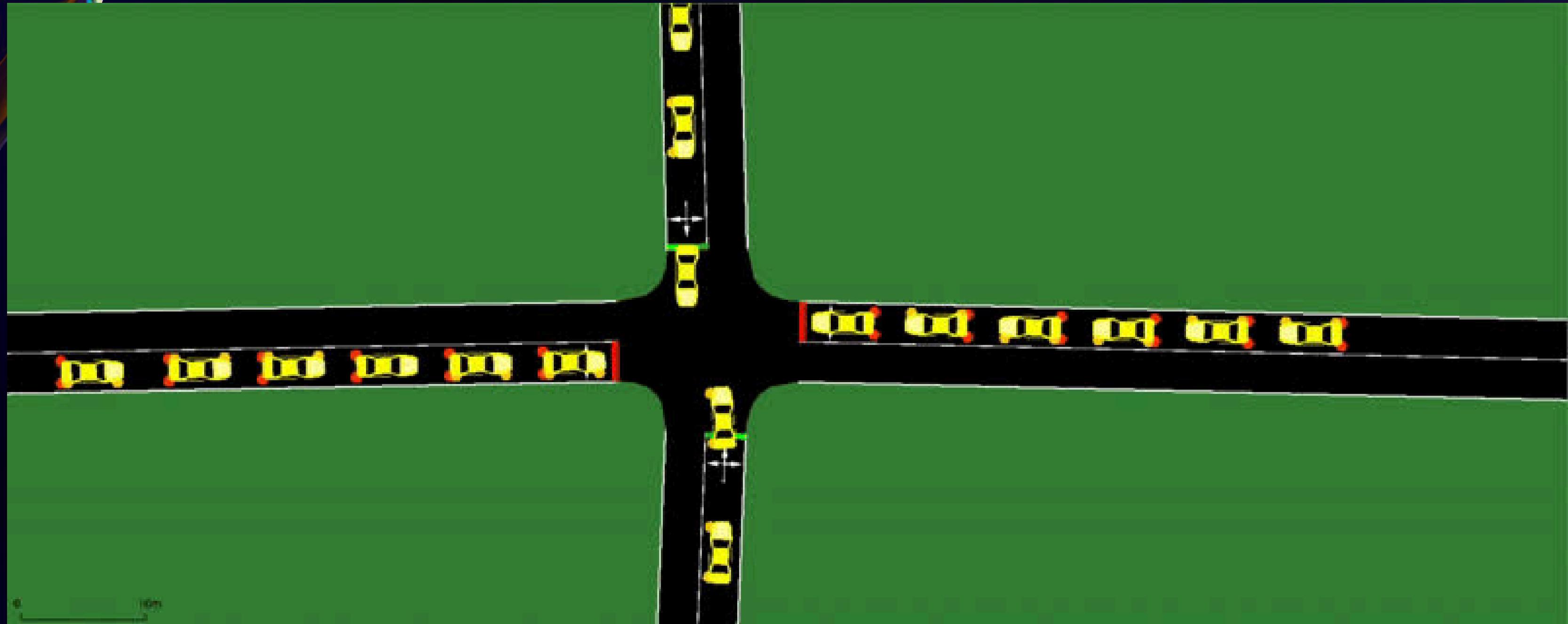
- 0: Move left
- 1: Move down
- 2: Move right
- 3: Move up

The **slippery mode** incorporates a $\frac{1}{3}$ probability for intended action. $\frac{2}{3}$ chances for slipping to perpendicular direction.

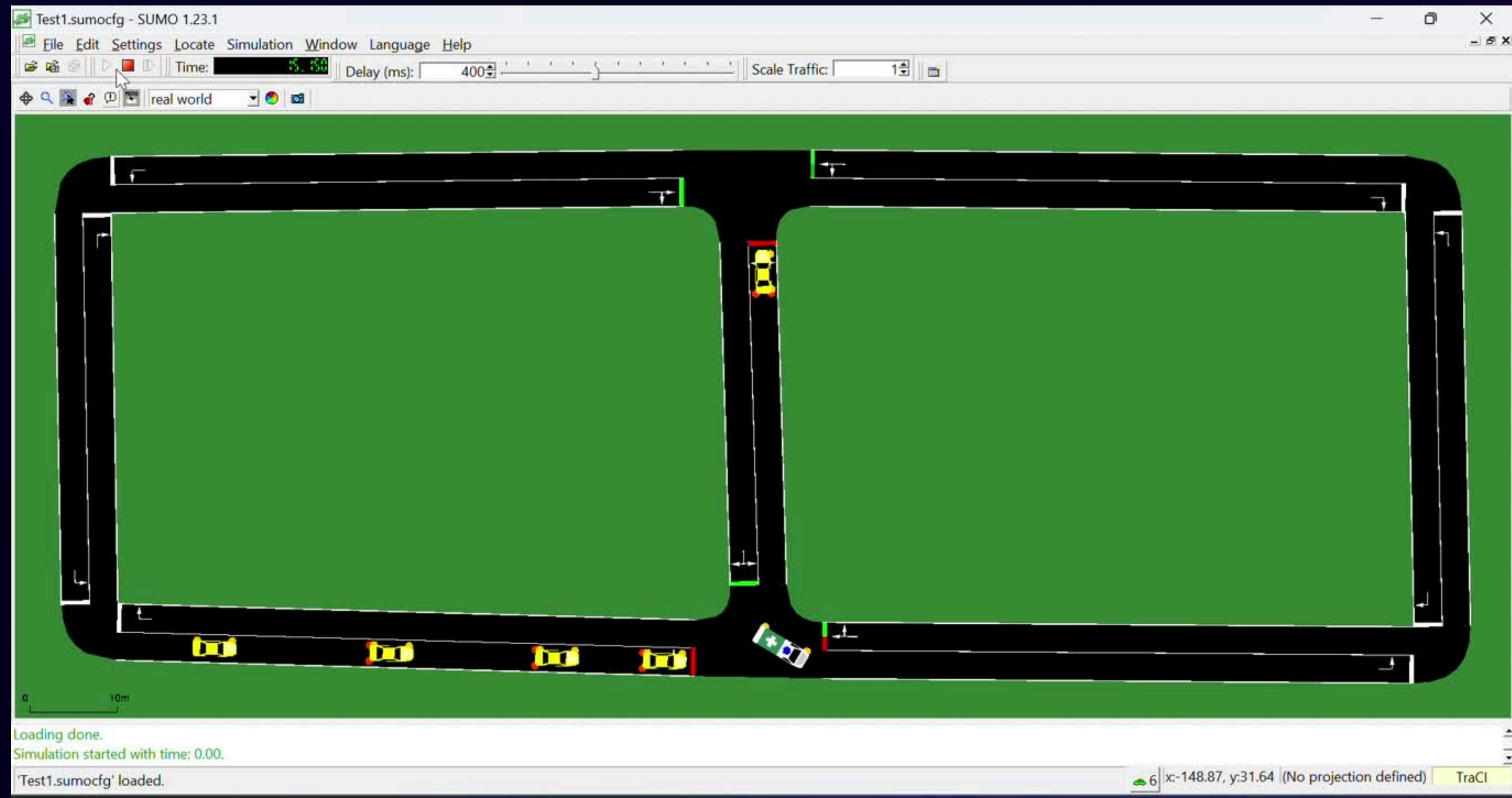
Performance Evaluation



Fixed TLS Logic in a 4-way intersection using SUMO



SUMO Traffic Light: Signal Preemption/Priority for Emergency Vehicles



Project Workflow

01

- Genetic algorithm for optimizing among many possible solutions.
- Q Learning & DQN on a single intersection

02

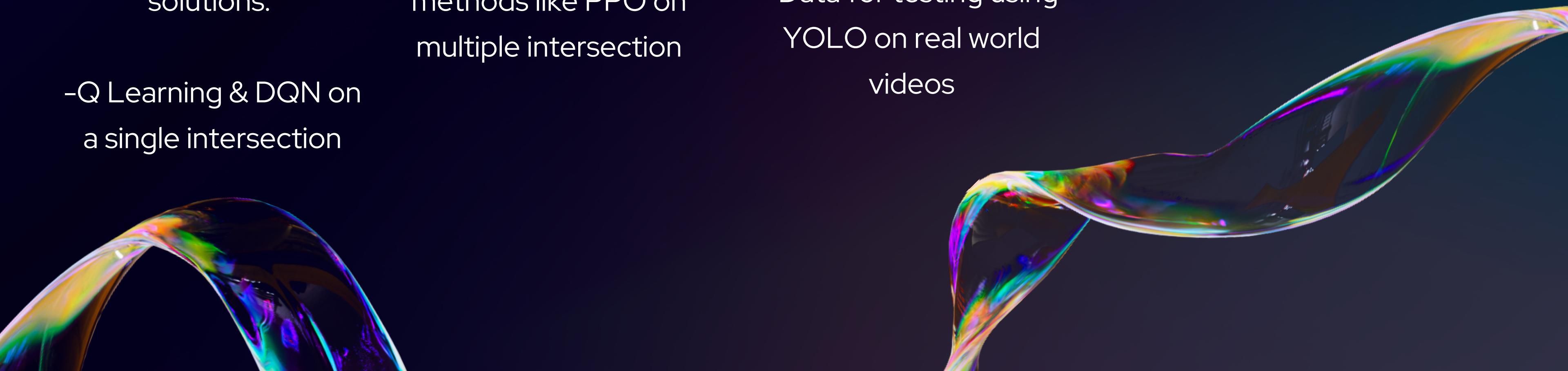
Multi Agent implemented with Policy Gradient methods like PPO on multiple intersection

03

- Implementing real city maps using OSM Web Wizard
- Data for testing using YOLO on real world videos

04

Performance Evaluation using RESCO Metrics





Thank You

Mentees: Ojas Alai, Sofia Abidi

Mentors: Yashvi Mehta , Mahi Palimkar