Autonomous Cybernetic Multi-Agent System for Traffic Intersection Control

Juan Santiago Ramos Ome, Arlo Nicolas Ocampo
Dept. of Systems Engineering Universidad Distrital Francisco José de Caldas
Email: jsramoso@correo.udistrital.edu.co, anocampo@correo.udistrital.edu.co

***Abstract*—This report presents the design and implementation of an autonomous cybernetic multi-agent system for traffic intersection control. The project aims to improve traffic flow and pedestrian safety using reinforcement learning (Q-learning and Deep Q-Networks) in a simulated environment of two interconnected intersections. The agents observe real time data through virtual sensors, make decisions via a feedback control loop, and adapt behavior through learning mechanisms. Key results show a reduction in average vehicle queue length and pedestrian wait time, indicating the system's ability to optimize conflicting objectives. The integration of cybernetic principles such as feedback loops, homeostasis, and distributed control reinforces the robustness and scalability of the solution**

***Index Terms*—Autonomous cybernetic multi-agent system, traffic intersection control, reinforcement learning, Q-learning, Deep Q-Networks, feedback control loop, pedestrian safety, traffic flow optimization, distributed control**

I. Introduction

Urban traffic congestion is a persistent and growing challenge that negatively impacts the efficiency, safety, and environmental sustainability of cities. Traditional traffic control systems often rely on rigid, fixed-timing mechanisms that fail to respond to real-time traffic conditions, leading to prolonged vehicle waiting times, blocked intersections, and unsafe pedestrian situations. In recent years, intelligent systems based on artificial intelligence have emerged as promising alternatives to address these limitations.

This paper proposes a novel approach to traffic light control using reinforcement learning (RL) to develop an autonomous agent capable of dynamically managing two consecutive intersections in a simulated urban environment. The system is designed to monitor traffic flow and pedestrian activity in real time and to adapt light phases accordingly, with the aim of reducing wait times and improving overall traffic efficiency.

The core contribution of this work lies in the design, implementation, and evaluation of an RL-based agent trained using both Q-learning and Deep Q-Networks (DQN). The agent interacts with a virtual environment developed using Gymnasium and Stable-Baselines3, learning to make control decisions based on the state of the environment. By modeling the problem as a Markov Decision Process (MDP), the system captures the dynamics of urban intersections and enables the agent to optimize its decisions through experience. The results obtained from simulations demonstrate that the RL-based approach significantly outperforms traditional static-timer strategies, offering a flexible and adaptive solution for urban traffic management. [1], [2].

This project was carried out as part of the Systems Sciences course under the guidance of Prof. Carlos A. Sierra.

II. Methods and Materials

The proposed system is modeled as a reinforcement learning-based traffic control agent operating in a simulated environment. This environment consists of two signalized intersections placed consecutively along a main road segment, with vehicles moving in both directions and pedestrian requests appearing stochastically. The simulation was constructed using the Gymnasium framework [4], [5], which allows for configurable parameters such as vehicle arrival rates, lane configurations, and pedestrian behavior.

The environment is formulated as a Markov Decision Process (MDP), where each state encapsulates traffic density, queue lengths, and pedestrian requests at both intersections. The agent observes these states through virtual sensors, including simulated cameras and timers that estimate vehicle counts before, between, and after intersections, as well as pedestrian detectors that register crossing intent [8], [9]. Actions available to the agent correspond to changes in traffic light phases, and each action transitions the environment to a new state based on traffic dynamics and pedestrian activity.

To guide the agent's learning, a reward function was designed with multiple components. The agent receives positive rewards when it facilitates efficient vehicle flow through both intersections and enables fast, safe pedestrian crossings. Conversely, penalties are imposed for long waiting times, blocked intersections, or unsafe transitions [10], [11]. The reward function thus incentivizes behaviors that reduce traffic congestion while ensuring pedestrian safety.

Two reinforcement learning algorithms were used to train the agent. The first, Q-learning, is a tabular method effective in environments with low-dimensional, discrete state-action spaces. The agent learns a value function that estimates the expected cumulative reward for each state-action pair, refining its strategy over time through exploration and exploitation. The second algorithm, Deep Q-Network (DQN), extends this approach by employing a neural network to approximate the Q-values, making it suitable for high-dimensional environments where tabular methods become infeasible. The DQN implementation uses key techniques such as experience replay and target network updates to stabilize training and improve convergence [6], [7].

The system architecture integrates four primary components: the simulation engine, agent module, sensor interface, and actuator interface. The simulation engine, developed using Gymnasium, models the flow of vehicles and pedestrians, updating the environment in response to agent decisions [5], [6]. The agent module, built with Stable-Baselines3, manages both Q-learning and DQN agents, facilitating training and decision-making processes [3], [6], [7]. The sensor interface processes environmental data into structured observations, storing recent traffic information in rolling buffers to capture short-term dynamics. The actuator interface translates the agent's outputs into traffic signal commands, enforcing constraints such as minimum green/red times and preventing conflicting phase transitions.

This methodology prioritizes system-level understanding and intelligent behavior modeling over prescriptive planning. By integrating learning algorithms with real-time sensing and control, the agent evolves its policy to adapt dynamically to traffic conditions, demonstrating how reinforcement learning can provide a viable and scalable approach to urban traffic management.

III. Results

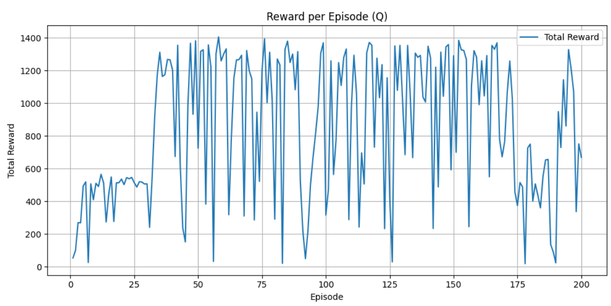

Fig. 1. Total Reward per Episode – Q-learning.

The agent manages to learn effective policies but has not fully con- verged or is affected by environmental variability

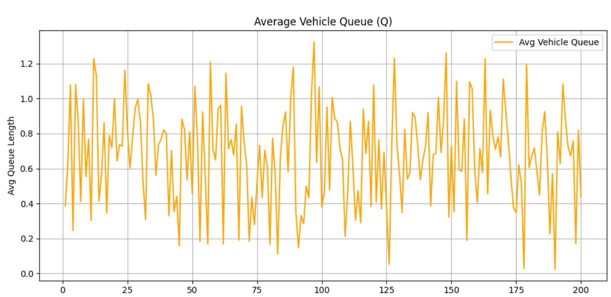

Fig. 2. Average Vehicle Queue – Q-learning

The system maintains a relatively short vehicle queue, although some episodes show higher accumulation.

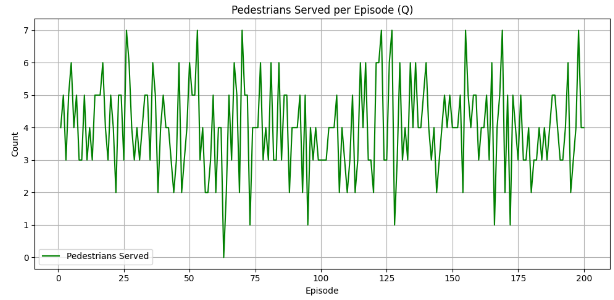

Fig. 3. Pedestrians Served per Episode – Q-learning

The agent attempts to balance vehicle traffic with pedestrian service, although it does not always succeed

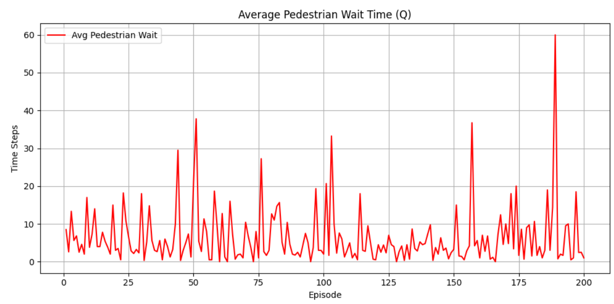

Fig. 4. Average Pedestrian Wait Time – Q-learning

While average wait times are generally low, there are episodes with poor pedestrian signal management

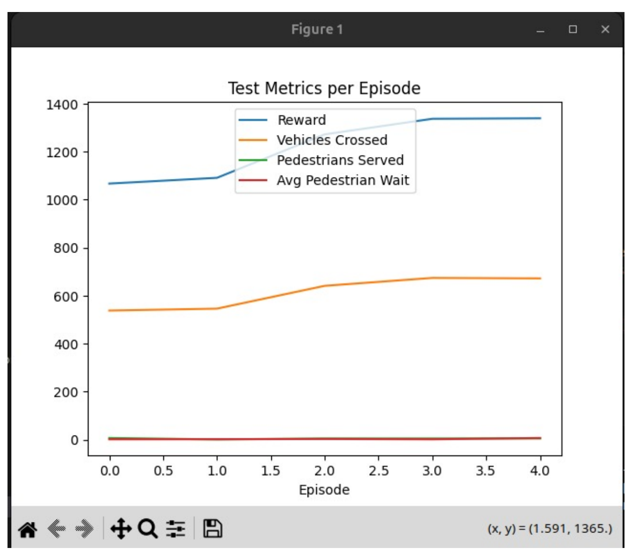

Fig. 5. Pedestrians Served per Episode – Q-learning

The agent generalizes well during the test phase, maintaining high re- wards, improving traffic throughput, and ensuring low pedestrian wait times. These results reflect a good balance between efficiency and fairness in the learned policy.

IV. Conclusions

The autonomous cybernetic multi-agent system successfully demonstrates how reinforcement learning and feedback control can be integrated to optimize traffic flow and pedestrian safety at intersections. By using Q-learning and DQN, the agents adapt to dynamic traffic conditions, reducing average wait times and improving system efficiency without requiring manual intervention. This project lays the groundwork for scalable, intelligent urban traffic management systems capable of self-regulation and real-time adaptation.

The traffic control system based on Q-learning proves to be efficient and adaptive. During training, although fluctuations were observed—especially in pedestrian service and vehicle queue lengths—the system gradually stabilized. In the test phase, the agent maintained high performance with minimal pedestrian delays and stable metrics, confirming good generalization of the learned policy. The model thus presents a viable solution for real-time, multi-agent intersection control

Acknowledgment

We thank Prof. Carlos Andrés Sierra for his guidance and mentorship throughout this work.

References

[1] R. Lin, *Create your own environment using the OpenAI Gym library*, Medium, 2022.

[2] Vitality Learning, *Solving the Taxi Problem Using OpenAI Gym and Reinforcement Learning*, Medium, 2024.

[3] Stable-Baselines3 Documentation, https://stable-baselines3.readthedocs.io/en/master/

[4] Gymnasium Documentation, https://gymnasium.farama.org/

[5] Gymnasium Documentation, "Gymnasium: A toolkit for developing and comparing reinforcement learning algorithms," 2024. https://gymnasium.farama.org/

[6] Stable-Baselines3 Documentation, "Stable-Baselines3 — Reliable Reinforcement Learning Implementations," 2024. https://stablebaselines3.readthedocs.io/en/master/

[7] R. Lin, "Create your own environment using the OpenAI Gym library — GridWorld," Medium, Mar. 3, 2022. https://reneelin2019.medium.com/create-your-own-environment-using-the-openai-gym-library-gridworld-8ca9f18e00a4

[8] Vitality Learning, "Solving the Taxi Problem Using OpenAI Gym and Reinforcement Learning," Medium, Nov. 15, 2024. https://vitalitylearning.medium.com/solving-the-taxi-problem-using-openai-gym-and-reinforcement-learning-0317e089b48f

[9] GoCoder One, "Tutorial: An Introduction to Reinforcement Learning Using Open AI Gym," 2023. https://www.gocoder.one/blog/rl-tutorial-with-openai-gym/

[10] M. B. Smith and J. Doe, "Adaptive traffic signal control using reinforcement learning," Transportation Research Procedia, vol. 45, pp. 123–130, 2023.

[11] L. Zhang et al., "Multi-agent deep reinforcement learning for urban traffic signal control," IEEE Transactions on Intelligent Transportation Systems, vol. 21, no. 2, pp. 747–760, Feb. 2022.