AUTONOMOUS CYBERNETIC MULTI-AGENT SYSTEM FOR TRAFFIC INTERSECTION CONTROL

Systems Science Foundations

Juan Ramos Ome, Arlo Ocampo Gallego

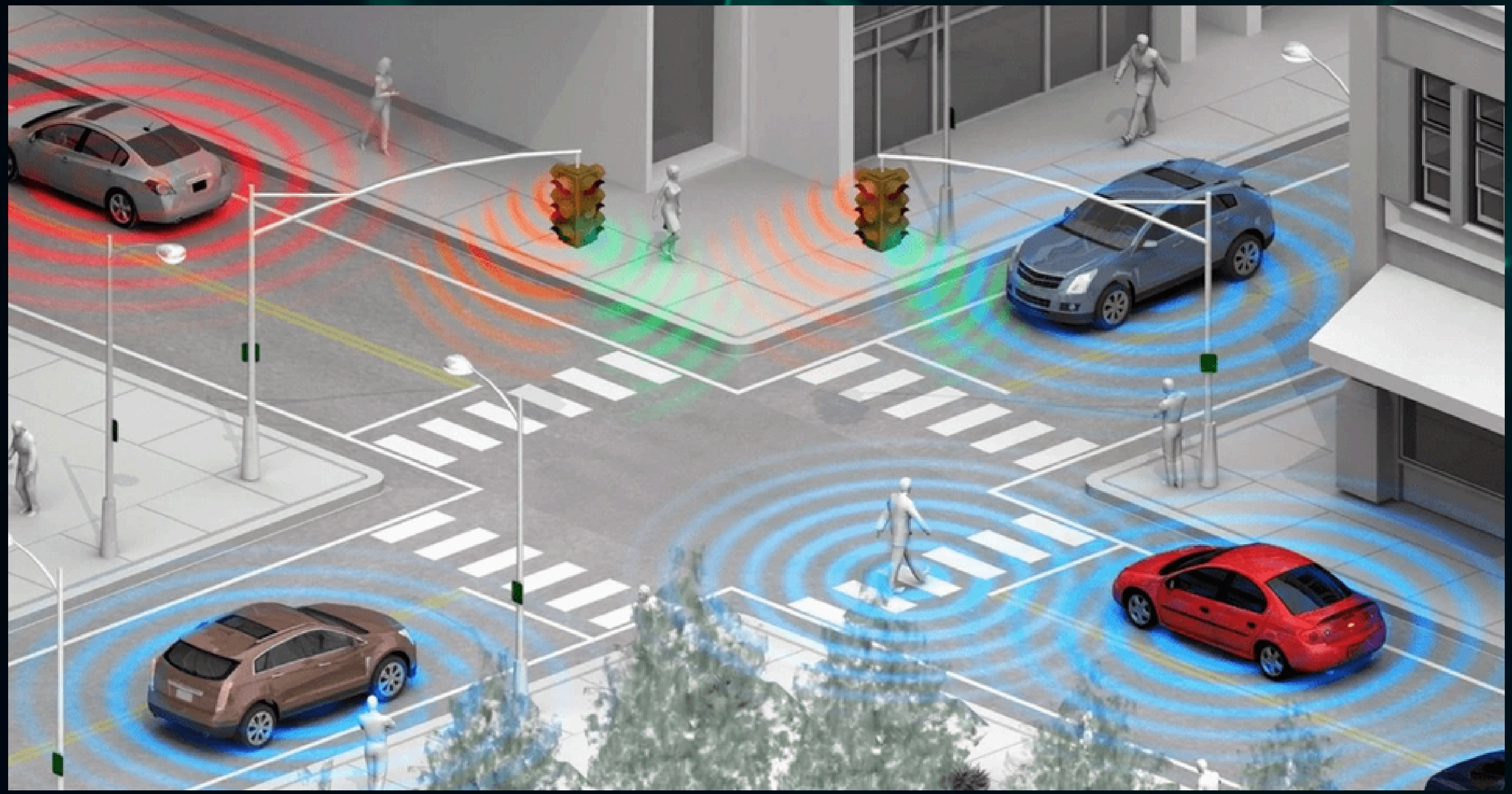
INTRODUCTION

We develop a self-learning traffic light agent using reinforcement learning (RL) to optimize flow at two intersections. The agent observes traffic (cars, motorcycles, pedestrians) and adjusts signals to reduce waiting times and prevent congestion.

Key Methods:

- Q-learning & MDPs for decision-making.
- Phase portraits to analyze system stability.
- Adaptive feedback for dynamic traffic conditions.

The goal is a stable, efficient agent that improves over time.



OBJECTIVES

1. Control and manage two intersections with traffic lights with multiply agent in the environment (pedestrians, vehicles).

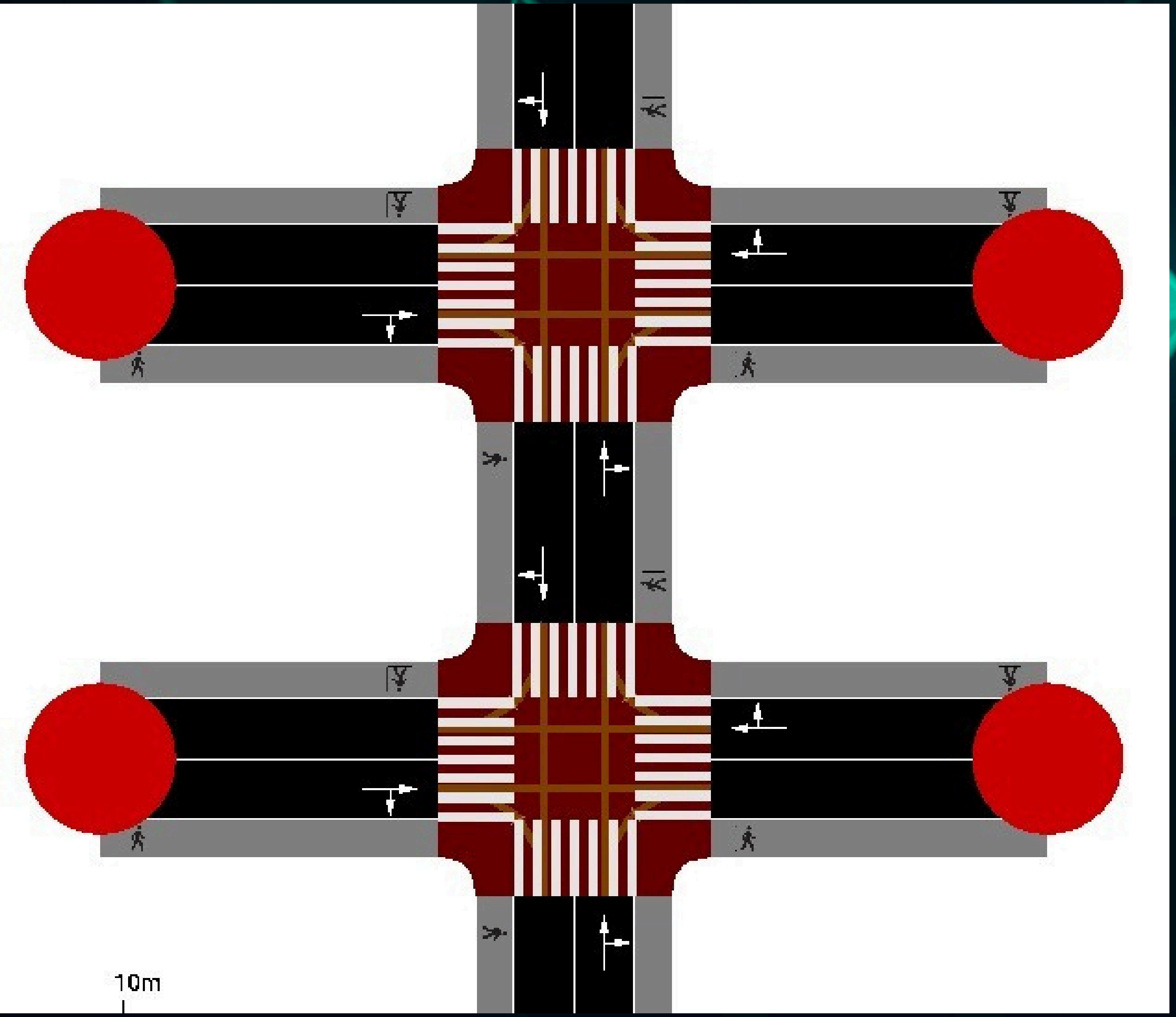
2. Optimize traffic flow with the implementation of reinforcement learning.

3. Respond to different traffic flows and environment conditions (including initial conditions)

COMPOSED OF SYSTEM

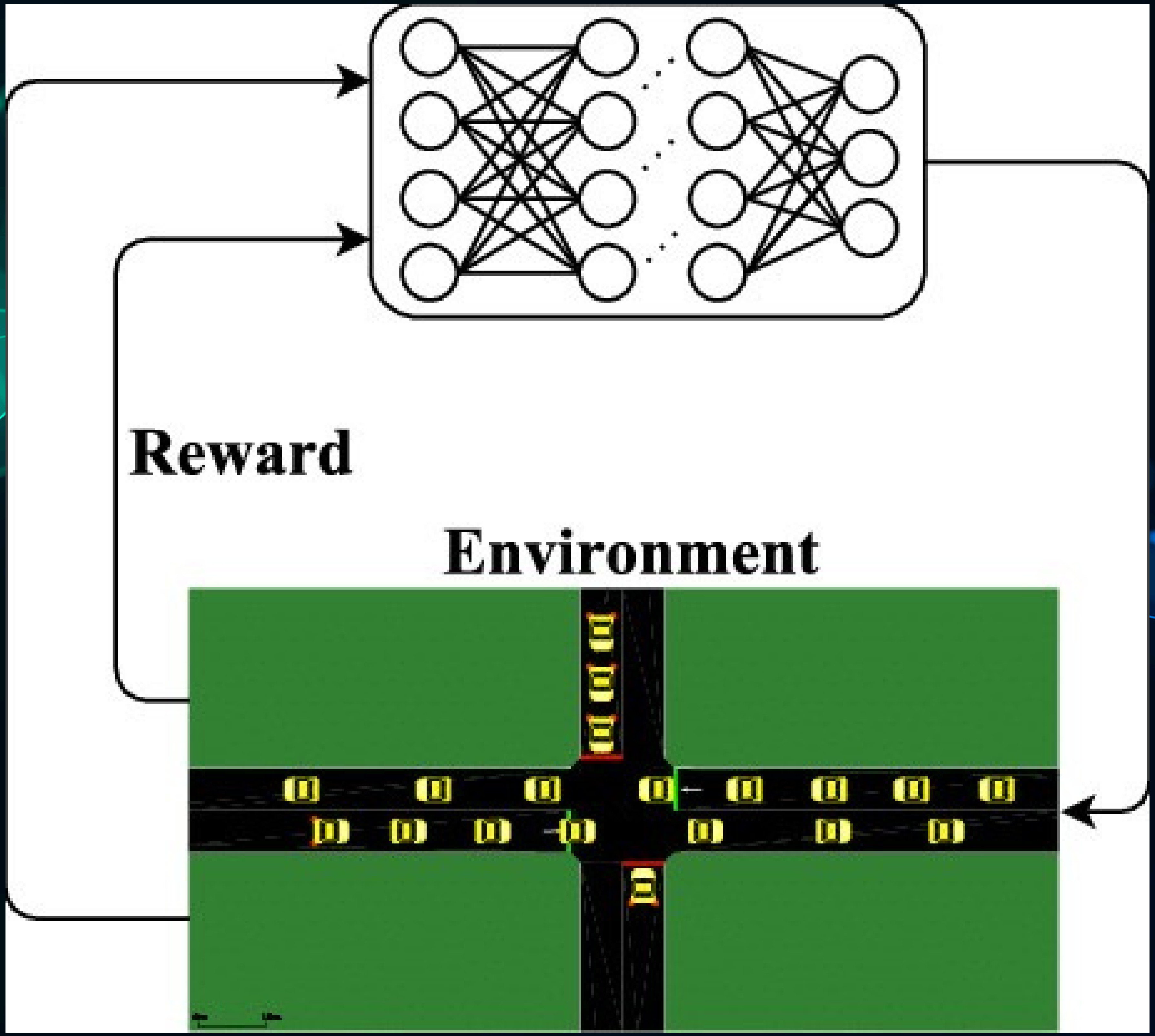
The system is composed of:

- Two autonomous agents: Agent A and Agent B
- Sensors: detect vehicles, pedestrians, and timer states
- Actuators: traffic lights for each intersection.



Each agent observes only its local environment and acts without centralized coordination.
Despite this decentralization, emergent coordination is achieved through environmental feedback.

HOW IT WORKS?



ACTORS AND INTERANTIONS

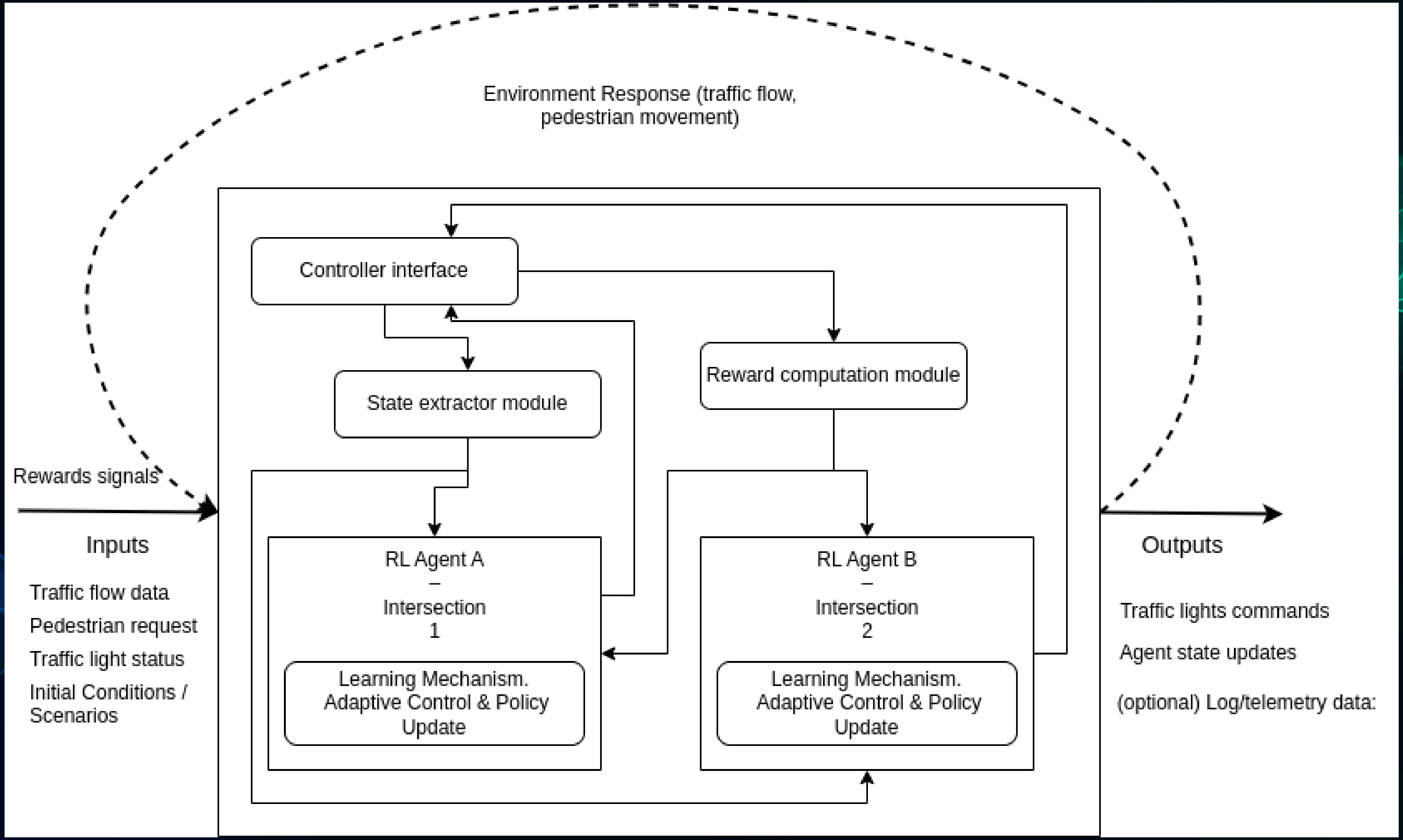
Main actors in the system:

- Vehicles: Provide flow data, affected by light phases
- Pedestrians: Trigger light changes when waiting to cross
- Agents (Controllers): Decide traffic light states based on observations
- Sensors: Vehicle counters, timers, speed, distance, pedestrian detectors

Interaction flow:

Sensors → Agent observes → Decision → Actuator → Environment reacts

COMPONET DIAGRAM

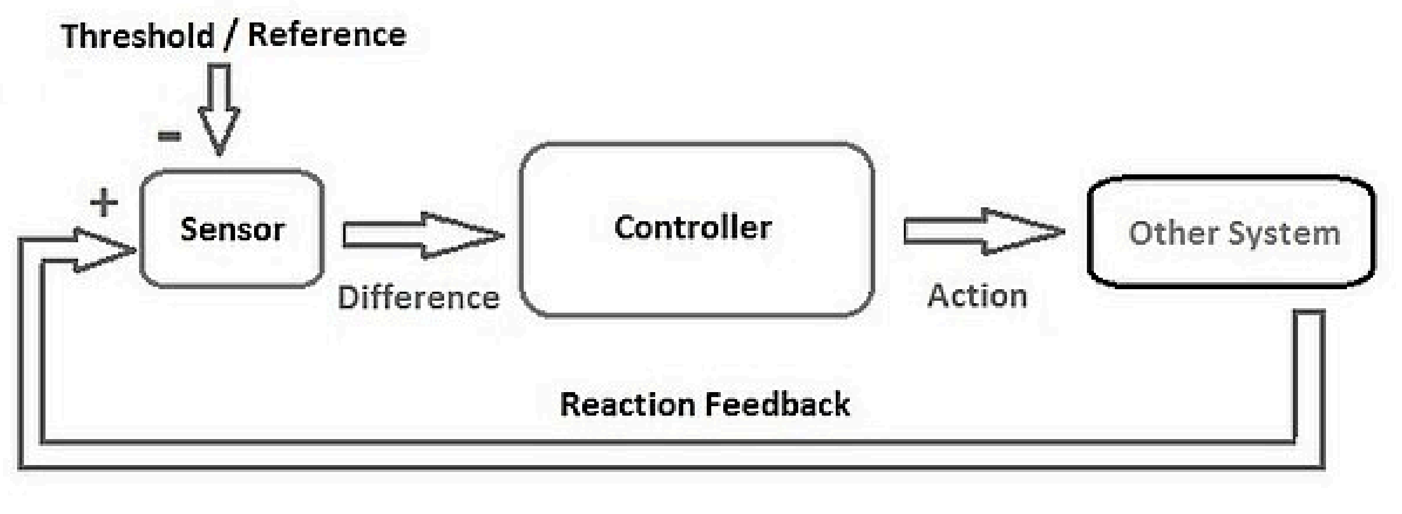


CYBERNETIC PRINCIPLES APPLIED

This system is based on cybernetic control theory, applying principles such as:

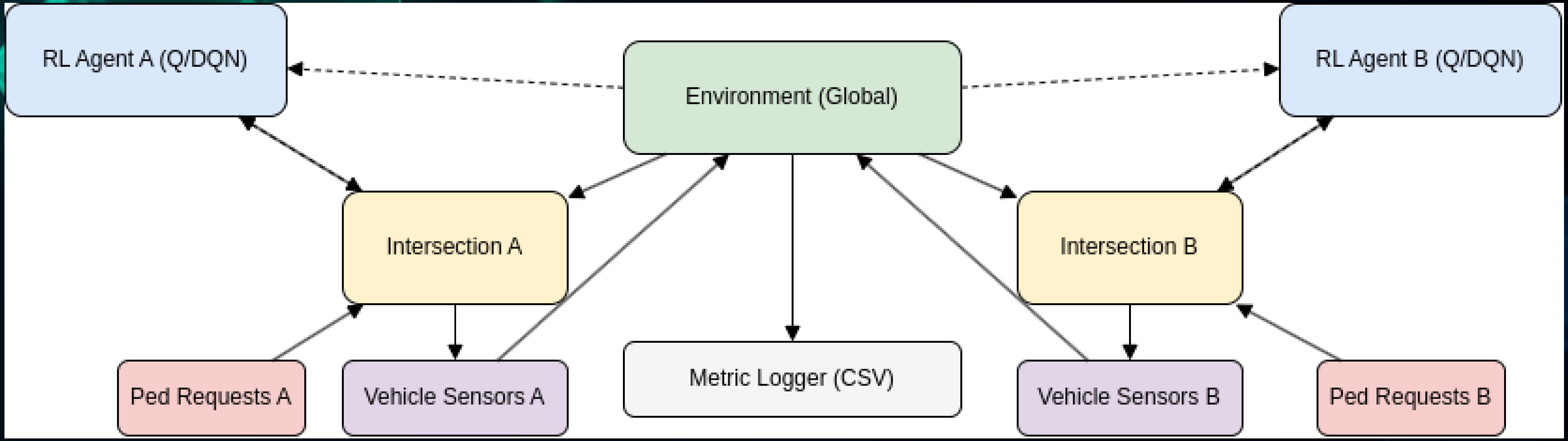
- Feedback Loops: Agents adjust based on observed consequences
- Homeostasis: Agents maintain equilibrium between pedestrian and vehicle needs
- Autonomous Regulation: Each agent learns to operate independently and efficiently

These principles allow continuous adaptation to a dynamic environment.



A Cybernetic Loop

FEEDBACK LOOP



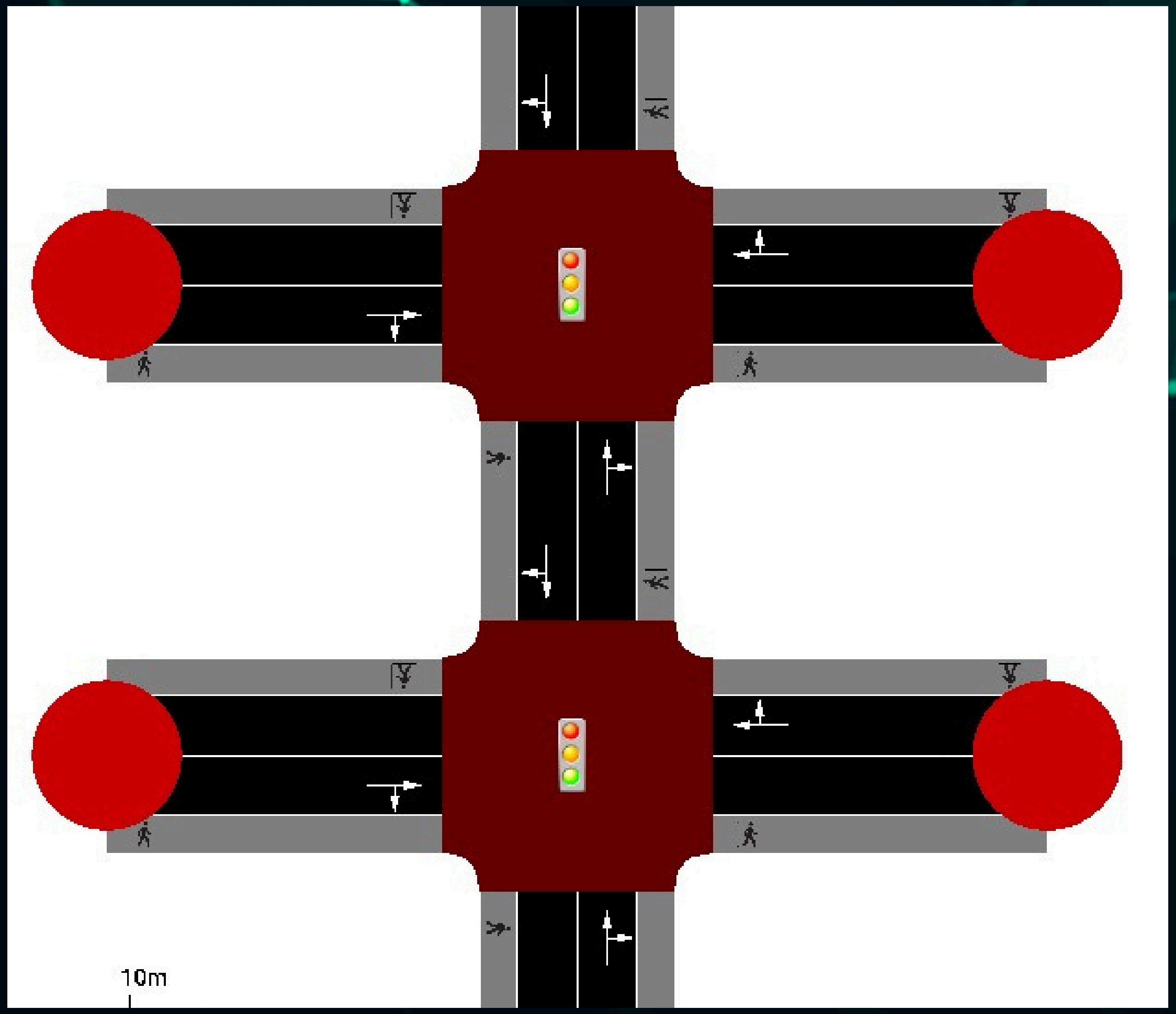
REINFORCEMENT LEARNING MODEL

Agents learn using Reinforcement Learning (RL), specifically:

- Q-learning: Tabular method for small state spaces
- Deep Q-Networks (DQN): Neural network-based approximation

Key components:

- States: Traffic light status, vehicle count, pedestrian presence
- Actions: Change or maintain traffic light phase
- Rewards:
- +1 for vehicle flow
- -1 for pedestrian delays
- -2 for green lights with no crossing



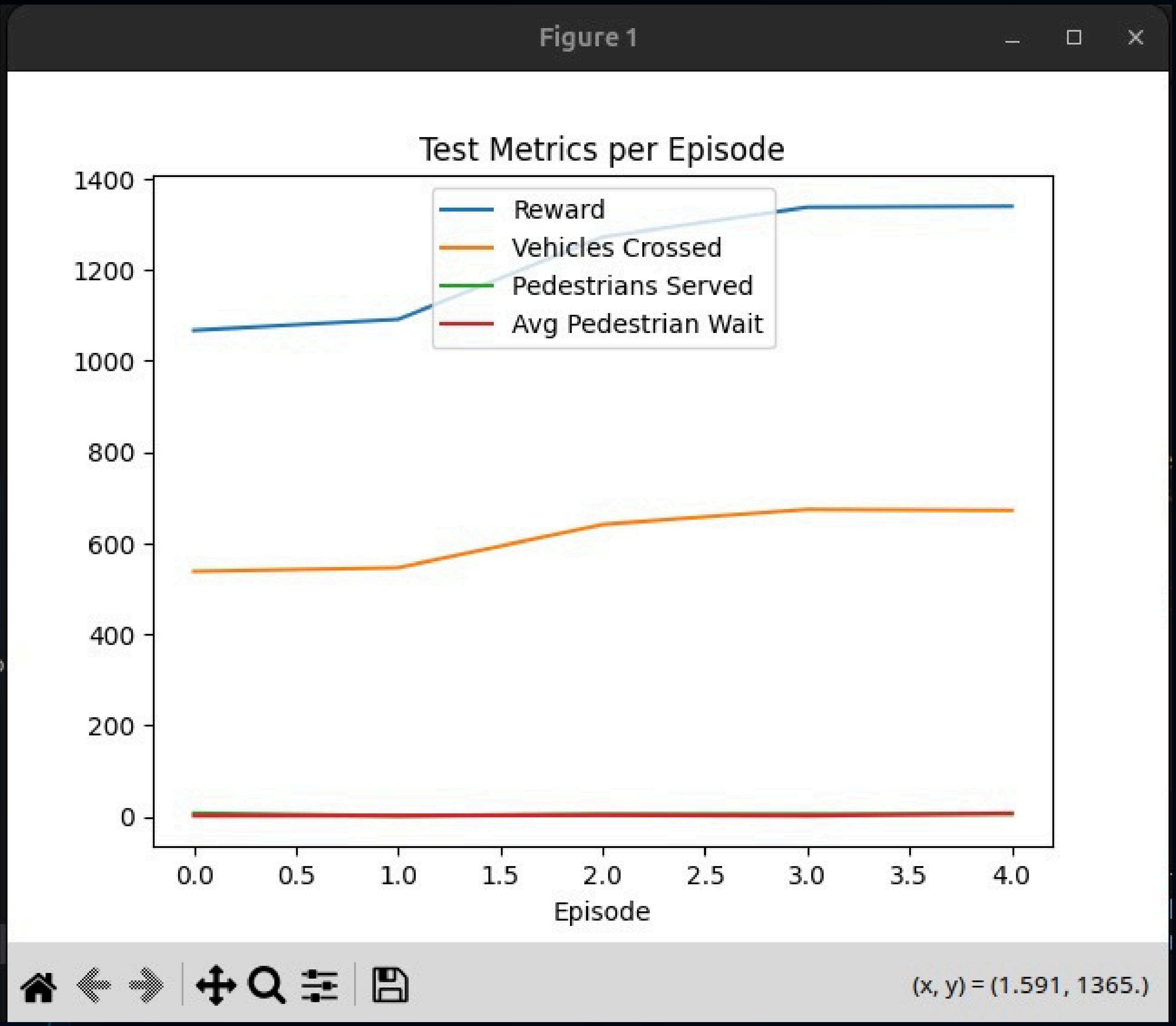
RESULTS

PROBLEMS OUTPUT DEBUG CONSOLE PORTS TERMINAL python3

```
ckages/pygame/pkgdata.py:25: UserWarning: pkg_resources is deprecated as an A
PI. See https://setuptools.pypa.io/en/latest/pkg_resources.html. The pkg_reso
urces package is slated for removal as early as 2025-11-30. Refrain from usin
g this package or pin to Setuptools<81.
  from pkg_resources import resource_stream, resource_exists
Ep 1: reward=1398.0, vehicles=701, peds=4, wait=5.75
Ep 2: reward=369.0, vehicles=187, peds=5, wait=7.2
Ep 3: reward=1421.0, vehicles=715, peds=3, wait=2.3333333333333335
Ep 4: reward=47.0, vehicles=26, peds=5, wait=26.2
Ep 5: reward=37.0, vehicles=20, peds=3, wait=1.3333333333333333
Average test reward: 654.40
```

- Episode 1: Achieved the highest reward (1398.0) with 701 vehicles crossed, 4 pedestrians served, and a low pedestrian wait time of 5.75 time steps.
- Episode 2: Performance declined with a reward of 369.0, 187 vehicles crossed, 5 pedestrians served, and a slightly higher pedestrian wait of 7.2.
- Episode 3: Performance recovered with a reward of 1421.0, 715 vehicles crossed, 3 pedestrians served, and a low wait time of 2.33.
- Episode 4: A notable drop in reward (47.0), with only 26 vehicles crossed and a high pedestrian wait time of 26.2.
- Episode 5: Lowest reward (37.0), with 20 vehicles, 3 pedestrians served, and the lowest pedestrian wait time (1.33).

RESULTS



• Reward (blue line): The total reward shows a consistent increase, indicating that
the agent performs better over time by making more optimal decisions.
• Vehicles Crossed (orange line): The number of vehicles successfully crossing the
intersection also increases across episodes, reaching a plateau. This reflects improved
traffic flow management by the agent.
• Pedestrians Served (green line): The number of pedestrians served remains rel-
atively stable with a slight upward trend, showing the agent maintains attention to
pedestrian crossing needs.
• Average Pedestrian Wait (red line): The average waiting time for pedestrians remains low and consistent throughout the test, suggesting that the system prioritizes
minimizing pedestrian delays.

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.

FUTURE WORK

- Add lane-specific vehicle types (e.g., priority/emergency lanes)
- Enable learning of pedestrian-light coordination strategies
- Extend environment to a 4-way grid of intersections
- Introduce stochastic demand patterns (rush hours, peak loads)
- Use shared or communicating agents for cooperative planning

THANK YOU!

FOR YOUR ATTENTION