

# Precise Design of Research

**Date:** 11/05/2025

**Week:** 10

**Author:** Alvin B. Lin

**Student ID:** 112652040

## Discussion 1: Design of The Project

### Overview and Motivation

The long-term goal of this project is to explore how future AI systems may gain the capability to perform **real-time, city-scale adaptive traffic control** that optimises global traffic flow while reasoning about dynamic disruptions such as accidents, weather, and demand spikes. Such a capability is not currently attainable: today’s AI systems can optimise small, simulated traffic grids, but cannot jointly (1) understand the continuous physical flow of vehicles, (2) integrate probabilistic incident information, and (3) coordinate thousands of intersections in real time.

We aim to outline the conceptual and mathematical foundations of such a future system, identify the necessary ingredients, and design a small “solvable model problem” that serves as the first step toward this long-term vision. This report is therefore a **plan**, and no experiments or implementation are presented yet.

### Future AI Capability: What AI Cannot Do Today but May Do in 20 Years

The envisioned future AI system would perform **city-wide, physically grounded, self-learning traffic management**, meaning:

- It understands and predicts the continuous flow of vehicles on every road segment in a heterogeneous, realistic network.
- It integrates multimodal inputs (sensor feeds, GPS flows, historical jam data, weather patterns).
- It reasons about stochastic events such as accidents, congestion spikes, construction zones.
- It coordinates thousands of traffic lights as a coupled dynamical system rather than as independent controllers.

The system would select traffic signal timings and phases that globally minimise congestion, fuel consumption, and travel-time variance. The significance lies in its potential impact on economic productivity, carbon emissions, and quality of life in urban areas.

To describe the capability concretely:

AI goal:  $a_t = \pi^*(o_t)$  where  $a_t$  controls all intersections and  $o_t$  captures real-time physical state.

This is an ability that requires both deep physical reasoning and large-scale optimisation—something no current AI system can perform reliably.

## Ingredients Required to Realise This Ability

To reach this capability, several key components are needed. These are not just keywords but functional elements in a full system.

### 1. Data (Multimodal, Real-Time, Spatially Coupled)

- High-resolution traffic density data across the entire road network.
- Real-time vehicle trajectories from GPS, sensors, and camera systems.
- Probabilistic incident data describing likelihood of congestion or accidents.
- Environmental context such as weather, time-of-day patterns, and special events.

The AI must learn relationships between these data types and global traffic behaviour. This requires synchronised streams because traffic is a spatially propagating phenomenon.

### 2. Tools (Mathematical + Computational)

- **Partial differential equations** to model traffic density evolution:

$$\frac{\partial p}{\partial t} + \frac{\partial f(p)}{\partial x} = 0.$$

- **Graph neural networks** to propagate information across intersections.
- **Diffusion or generative models** to simulate plausible traffic disturbances.
- **Reinforcement signals** for decision-making under uncertainty.
- **Symbolic or physics-guided reasoning** to maintain consistency with flow conservation laws.

Each tool plays a distinct role: PDEs provide physical structure, GNNs encode spatial connectivity, RL provides sequential optimisation, and generative models simulate disturbances.

### 3. Hardware / Environment

- Real-time sensor networks across the city.
- Edge-computing infrastructure at intersections to process local flows.
- High-performance compute clusters to support global coordination.

A system of this scale requires distributed computation and continuous sensing.

### 4. Learning Setup

A hybrid learning framework is required:

- **Self-supervised learning** for traffic representation learning from unlabeled flows.
- **Meta-learning** to adapt quickly to new events (emergencies, construction).
- **Reinforcement learning** for long-horizon control.

- **RLHF or constraint-based fine-tuning** to satisfy safety and fairness requirements.

These components interact to form a scalable, robust decision-making pipeline.

### Machine Learning Types Involved

The envisioned system is inherently a **combination of methods** rather than a single paradigm.

- **Supervised learning** is used for predicting short-term flow patterns and incident probabilities:

Data input → Predicted density and congestion risk.

- **Unsupervised / self-supervised learning** discovers low-dimensional representations of traffic dynamics.
- **Reinforcement learning** is necessary for global control:
  - The environment is interactive.
  - The reward signal is derived from congestion reduction.
  - Actions shape future traffic flow.

There is no fixed “correct answer” to the control problem; only long-term performance matters, which makes RL essential.

### Solvable Model Problem: First Step Toward the Long-Term Goal

To explore the foundations of this future capability, we design a small, solvable model problem.

#### Problem Design



**Figure 1:** Road Arrangement in Barcelona, Spain(Source: <https://shorturl.at/5BObK>)

We simplify the full traffic system to a single ‘Tian’-grid composed of intersections, like the road in Barcelona (Figure 1). Let the density on each road be described by:

$$\frac{\partial p(x, t)}{\partial t} + \frac{\partial}{\partial x} \left( v_{\max} p(x, t) \left( 1 - \frac{p(x, t)}{p_{\max}} \right) \right) = 0.$$

The traffic light alternates between the two directions, and the AI must choose the switching time.

- **Input:** traffic densities on both roads.
- **Output:** the next switching time or duration.
- **Goal:** minimize queue length and total waiting time.
- **Data:** synthetic PDE-simulated traffic flows.

This is a minimal version of the full problem but retains physical structure.

#### **Model and Method (Planned)**

A simple reinforcement learning model such as Q-learning or a small policy network will be used:

$$a_t = \pi_\theta(o_t), \quad r_t = -(p_{\text{wait},1}(t) + p_{\text{wait},2}(t)).$$

This setup allows us to test whether a learner can pick switching times that reduce congestion in this toy environment.

#### **What We Plan to Learn**

- Whether the RL agent can discover non-trivial signal patterns from PDE-driven traffic.
- How sensitive the control performance is to noise or incident disturbances.
- How representations of traffic density influence learning efficiency.

The lessons from this toy model would indicate how far current methods are from achieving the long-term goal and where the main obstacles lie, such as:

- the difficulty of credit assignment over long time horizons,
- the need for spatial coordination,
- the gap between simulated and real-world complexity.