# Thesis Outline

Author: Brian King

Completion Deadline: March 2023

## Motivation

On roads in the United States, many drivers travel at velocities above the speed limit. On some stretches, this number can be higher than 70%[[1]](#footnote-1). When traversing a corridor of traffic lights, this can lead to extremely low fuel efficiency. While many drivers will slow down when approaching a red light to reach the intersection at a green light, selecting an appropriate speed is virtually impractical without knowledge of the traffic light’s real-time signal phasing and timing (SPaT) data. By allowing a computer to receive and analyze SPaT data from upcoming lights, target velocities can be selected that maximize fuel efficiency.

## Problem Statement

As vehicle-to-infrastructure technology (V2I) becomes more prevalent, fuel consumption can be decreased by commanding vehicles to drive at an optimal velocity through a stretch of V2I-enabled traffic lights (i.e., connected corridor). There is no consensus on the most optimal algorithm to select target velocities, so exploration into control methods such as reinforcement learning (RL) may lead to increased fuel efficiency on our trafficways.

## Objectives

The research objective is to develop a reinforcement learning agent that can be deployed on any connected corridor (without prior knowledge of the light positions) and make real-time decisions to decrease fuel consumption. This research will evaluate how individual vehicle and overall corridor fuel consumption is affected with the implementation of the RL controlled vehicle.

## Proposed Approach

### Task 1- Simulation Environment and Agent Feedback Creation

Using a combination of SUMO (Simulation of Urban MOBility) and Python, a traffic network will be created to simulate a V2I-enabled connected corridor of streetlights. This corridor will include multiple intersections with variable positions and signal phasing and timing (SPaT) data, as well as traffic flow. A Deep-Q Learning network (DQN) will be built in Python. The rewards and state observations from SUMO will be sent to an experience buffer which the DQN will randomly sample from to train the reinforcement learning agent as shown in Figure 1.

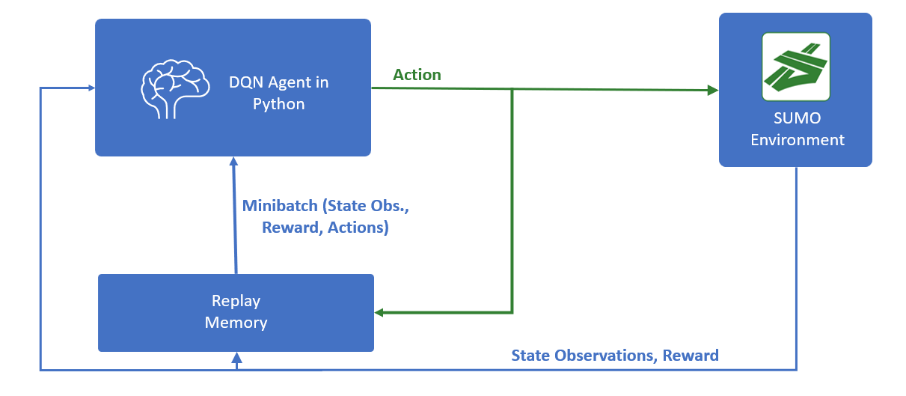


Figure . Program Architecture

The observations will include the current **vehicle velocity, the distance until the intersection(s), the current signal phases, the time until the phase change(s),** and **the initial phase time(s)**. Another observation that can be included is the **traffic density** on each stretch of road. This can be done using detectors/cameras at intersections, or with the use of historical data. Historical data allows for a broader application of the RL agent, as detector/cameras are not present at every intersection. The agent will select a target velocity which the vehicle in SUMO automatically attempts to hit using a pre-set acceleration. The agent will output a target velocity (action) at every intersection based on the state observation. Once the vehicle reaches the next intersection, the reward is sent back to the agent to update its policy (Figure 2). This will keep the agent from constantly updating and changing its target velocity, leading to more stability in the training in comparison to an agent that must make a decision at a specific time step of the simulation.

Timeline

Description automatically generated

Figure . DQN Agent Decision Timeline

### Task 2- Build scenarios:

1. *RL Agent vs rule-based fuel efficiency in traffic*

The RL agent will be analyzed in comparison to a baseline vehicle in SUMO. The baseline vehicle will attempt to hit a specific velocity using pre-determined acceleration values. The baseline vehicle does not receive SPaT data outside of the phase of the next signal.

1. *Observation horizon performance on RL Agent*

The agent will be trained with knowledge of SPaT data from different numbers of intersections. For instance, the RL agent can be trained with knowledge of only the SPaT data from the upcoming traffic signal. It will then be compared to an agent that was trained with knowledge of the next three upcoming traffic signals.

1. *RL Agent’s effects on fuel consumption of entire corridor*

While the agent should be able to decrease its fuel consumption, it may have a negative effect on fuel efficiency of surrounding vehicles. For example, if the RL agent commands a target velocity that is too small, subsequent vehicles may not be able to pass through an upcoming traffic light, leading to unnecessary stops and therefore higher fuel consumption.

1. *Fuel consumption study with multiple RL Agents*

A scenario where multiple RL agents exist will be analyzed to determine the performance of these vehicles when multiple are deployed. The agent will be trained with only baseline vehicles, so adding more RL agents may not necessarily increase fuel efficiency, as the RL agent may overfit to scenarios with only baseline vehicles.

1. *Training with multiple RL agents*

Due to the aforementioned overfitting issue, multiple DQN agents can be trained on the intersection at once. This interaction may decrease the overall fuel consumption for the entire intersection as more RL agents are deployed.

1. *Multiple lanes instead of single lane corridors*

### Task 3- Evaluation of scenarios

Each of the six scenarios will be analyzed by comparing fuel consumptions to baseline values. The speed limit, light positions, as well as signal timing can be altered to perform a sensitivity analysis for which parameters effect the agent most.

1. https://www.tandfonline.com/doi/pdf/10.1080/014416499295420?needAccess=true [↑](#footnote-ref-1)