ME 691: Advanced Machine Learning Project Proposal
Max Schrader
03/27/21

Project Title:

Using Reinforcement Learning to Optimize Traffic Signal Timing in SUMO

Background:

The DOE project is a joint venture between the University of Alabama and Oak Ridge National Laboratory, with the aim being improved network-wide fuel economy. Initial findings from a test corridor in Tuscaloosa show that improved light timings correspond to a 2-12% reduction in fuel consumption. As part of the project, a partnership with the German Aerospace Center (DLR) and specifically their SUMO traffic simulation software has been created, giving UA the tools to perform microscopic traffic simulations of the Tuscaloosa region and to test potential light timing improvements. SUMO has a built-in vehicle model, but the ability of that model to accurately capture vehicle dynamics is under question. To get a better estimate of energy efficiency improvements, Oak Ridge and UA are partnering to develop more detailed vehicle models.

The optimization of traffic light timing has been researched extensively in the past, but typically focuses on increasing vehicle average velocity, maximizing vehicle flow or decreasing waiting time. Most of the research in combining reinforcement learning (RL) and traffic simulation has been in influencing connected vehicle and autonomous vehicle (AV) control, with one of the classic examples being RL controlled AV's dissipating traffic waves and bottleneck decongestion.

My project will focus on combining the SUMO micro-simulation vehicle emissions models and traffic signal optimization, using RL to potentially achieve significant improvements in overall system energy efficiency.

Environment Model:

The environment model is a SUMO model of three intersections in Northport, Alabama. The three intersections serve industrial, commercial (Walmart, Lowes) and residential areas respectively. The traffic in the simulation has been carefully calibrated to match real-world traffic as close as possible. Calibration is done by aggregated detector counts (stored in a SQL server) and deriving traffic routes that satisfy the counts most closely. To benchmark the RL-controlled traffic lights, the simulation will be ran with the lights behaving exactly the way that they did on any given day.

The observation space for the RL agent is the count of vehicles in each lane approaching the intersection, with some maximum distance that the "sensor" can pick up cars at. The action space is placing synthetic detector calls at the intersection, which will move the traffic light to the desired phase. The reward will be some combination of throughput and –1 * fuel consumption.

Approach:

Flow will be used to train the RL-agent controlling the traffic light. Flow is written in Python, along with the SUMO API. Custom code will be written to create both the state and action space for the RL-agent. According to results in the Flow benchmarks, Evolutionary Strategies, a gradient-free method, will be used to train the traffic signal controller.