Traffic Light Control using Reinforcement Learning

Daniel Goldberg and Andrew Elstein

Department of Electrical & Systems Engineering, Washington University in St. Louis, St. Louis, MO

Background

Traffic Lights are an important part of our transportation system, controlling the roads and keeping drivers safe. Yet, when not controlled properly, or in times of large traffic, these traffic lights can add a large amount of unnecessary waiting and traveling times for commuters, wasting time and resources. The goal of this project is to study, implement, and analyze algorithms used to control traffic lights. We used an algorithm of reinforcement learning to locally optimize each intersection of traffic lights by calculating the probabilistic waiting time of each car currently at the intersection.

Simulation

Structures

In order to test the traffic light control algorithms, we needed to design a simulation of automobile traffic. Using MATLAB, we created the simulation using the three structures: Cars, Roads, and Intersections. The Cars structure contains all the cars and their properties: position, maximum velocity, destination, route, color, etc.

The Roads structure stores the locations of all the cars on the road. The Roads structure is made up of two lanes and each lane stores a vector of 10 cells, which act as locations on the road.

The Intersections structure contains the status of the traffic lights, as well as which roads intersect at the given intersection.

Dynamics

The simulation mirrors the dynamics of real traffic. Yet, in order to make the simulation easier, we used discrete time and locations. In our simulation, no two cars can exist in the same cell, or position, on the road, as well as cars can not go through red lights, pass other cars on a one lane street, or collide into other vehicles. Once the simulation was completed, we could test our algorithms in our simulated traffic environment.

Visualizations

Each time an intersection is reached, the algorithm calculates the cumulative waiting time for each car at the given intersection, by calculating the probability of a car's position in the future given the current traffic light is either red or green for each road.

This figure shows the average traveling time of various systems of 1 to 40 cars, using both non-adaptive and adaptive traffic light controllers. The adaptive controller outperforms both the non-adaptive controllers for our system.

Algorithm

The reinforcement algorithm we implemented was from Marco Weiring's "Multi-Agent Reinforcement Learning for Traffic Light Control" [1].

The algorithm calculates the cumulative waiting time for each car at a given intersection, by calculating the probability of a car's position in the future given the current traffic light is either red or green for each road.

\[
Q((t_i, p, d), t) = \sum_{(t', p', d') \in T} P((t_i, p, d), (t', p', d')) \cdot (R((t_i, p, d), (t', p', d')) + \gamma Q((t', p', d', t + 1))
\]

This V (value) function gives a weighting of the relative benefit of a traffic light's decision for each car.

\[
V((t_i, p, d)) = \sum_{L \subseteq D} P((t_i, p, d), L) Q((t_i, p, d), L)
\]

The Reward function is described by a given reward of 0 for movement and 1 for staying still.

\[
R((t_i, p), (t', p')) = \begin{cases} 1 & \text{if } (t_i, p) = (t', p') \\ 0 & \text{if } (t_i, p) \neq (t', p') \\ \end{cases}
\]

The goal of the algorithm is to minimize the reward function, and through this it will learn to minimize the amount of time cars are stuck in a standstill.

By using these functions to determine the future probabilistic waiting times of all the cars at each intersection, the traffic lights can choose the light scenario which produces the least waiting time for the cars at the intersection.

Conclusions

In conclusion, the adaptive controller outperforms the 1-second non-adaptive controller by 15% in a sparse system (i.e. 5 cars), and performs the 1-second non-adaptive controller by 3% in a highly saturated system (i.e. ~40 cars).

Implementing this type of adaptive traffic light control could decrease average traveling times of drivers for systems of all sizes, but decrease the traveling time the most in systems with few cars.

References


Acknowledgements: Thank you to Dr. Ayed Nehorai and our mentor, Zhao Tan, for their help and support on our research.

Fig. 1: Visualization of a 10 car system.

Fig. 2: Average Traveling Time of Various Systems

Fig. 3: Total Waiting Times for High Traffic System