# Designing an optimum traffic signal system using reinforcement learning

Vaishakhi Mayya

December 2015

## 1 Introduction

Most large cities have to deal with serious traffic jams, especially in certain times in the day. These jams waste time for the people stuck in them, as they can't use that time for any productive activity. They are also harmful to the environment, increase emissions, fuel use and generally increase road-rage. Often, especially large urban centres, a major source for traffic jams are large intersections that are controlled by automatic traffic signals. The signals regulate the flow of traffic by ensuring that there are no two conflicting directions for the cars within the intersections. A poorly designed traffic signal system for a network of roads can act as a bottle neck causing large back-ups fairly quickly.

The aim of this project is to use reinforcement learning to model an optimum signal pattern for the traffic lights for a road and traffic flow. The agent learns the optimum pattern and is rewarded for minimizing the expected time on the designed road network for each car. it will be compared it against the expected time in the network with fixed time signals.

## 2 Reinforcement Learning

Reinforcement learning [1]

- uses positive/ negative reinforcement to train the agent
- specifies the actions, the agent can find the best possible action
- develop an ideal policy that maps states to actions

The agent can learn by observing the environment. The problem of minimizing traffic at a signal can be modeled as an RL problem. Here the signal states can be changed based on the average time spent by vehicles on the road. The agent can be trained to change the signals to minimize the average time spent in the system

### 2.1 Sarsa

SARSA or (state, action, reward, state, action) is an On- policy algorithm for learning markov decision processes. It is an iterative algorithm where at each step the Q-values for each state action pair are maximized. At each step, the agent executes the action, and obtains the reward for the action. each action pushes the agent to the next state, and the Q(s,a) for the previous state is updated using the following equation.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)]$$

where $\alpha$ is the learning rate and $\gamma$ is the discount factor to account for future rewards.

## 3 Traffic Model

The road model looked like the image shown below, with two signals controlling each road separately as shown in the figure 1. The vehicles were modeled as follows:

- At any time instant, a car would appear either on the vertical road or the horizontal road.
- The traffic load on the roads can be equal or more on a main road.
- If the car is before a signal and the signal is green, the vehicle can move forward, provided there is no one ahead of the them
- Once it crosses the signal it can move forward provided the space is empty

To model the signal behavior:

- Each road is controlled by a signal with two states, r or g.
- Since signals have to be disjoint, there are two states for the system - (r,g) and (g,r).
- The signals can stay in the same state or flip to the next state,and this was controlled by actions (stay, flip).

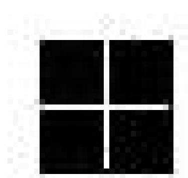

Figure 1: Intersection for the traffic model

# 4 Observations

## 4.1 Equal Traffic Loads

The following images show the effect of using salsa to train an agent to control the traffic signal in case of equal load, which means that the traffic was evenly distributed on both roads. This was compared against the current system with fixed time intervals for each signal. The time interval for the fixed time signal was set at 45 time units for each side. The figure 2 shows the average time (over 50 trials) spent by each individual car on the road. The x-axis denotes the vehicle id.

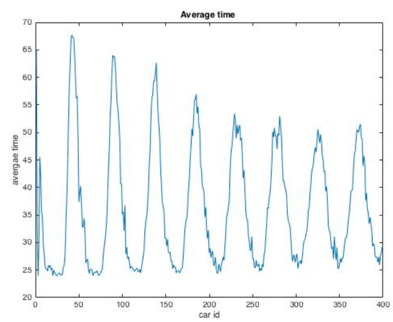


Figure 2: Average waiting time in a standard signal for 400 cars with equal load

Figure 3 shows the effect of using reinforcement learning for equal traffic . However, it can be seen that this does not show a significant improvement in wait times compared to using a fixed time signal

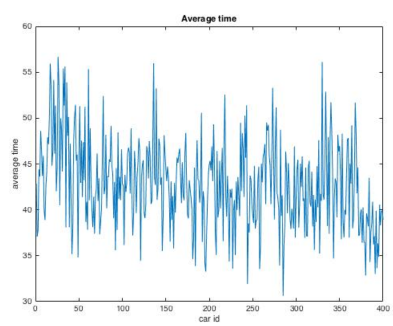


Figure 3: Average waiting time in the signal controlled by agent for 400 cars with equal load

## 4.2 Unequal Traffic Loads

The following images show the outcomes with unequal vehicle loads. About 90% of the traffic is coming in through the main road, (the horizontal road). And the remaining traffic comes through the other road. since the signal wait time is long, there is no difference in the average wait time for the fixed signal (figure 4).

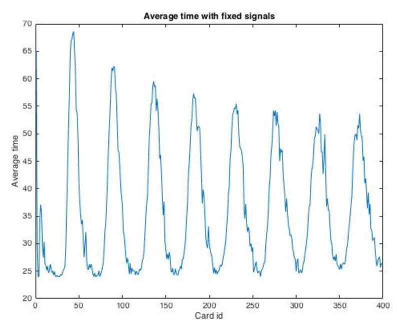


Figure 4: Average waiting time in a standard signal for 400 cars with unequal load

However with reinforcement learning, there is an improvement as the average time for most cars drop to under 40 time units as shown in the figure 5. The first few cars in the system show worse times but as the agent learns, the time improves.

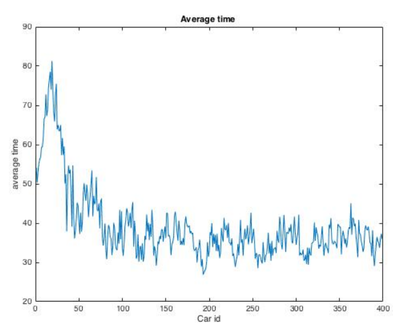


Figure 5: Average waiting time in the signal controlled by agent for 400 cars with unequal load

## 5 Conclusion

It is possible to use reinforcement learning to improve wait times at a signal for cars when there is heavier traffic load on one the roads. Here, the agent works by allowing the cars on the heavier traffic road have greater priority, thus reducing the wait time by hurting the people waiting at the other road. In the case of equal traffic on both roads, the using an trained agent does not produce a better outcome than using the conventional system. For future work, more signals can be added to create a network, so as to reduce the average time over the entire road network.

## References

[1] Barto A. Sutton, R. *Reinforcement Learning:An Introduction.* MIT Press, 1998.