In your report, mention what you see in the agent's behavior. Does it eventually make it to the target location?

The car wanders around aimlessly. It does occasionally make it to the end, but usually does not.

Justify why you picked these set of states, and how they model the agent and its environment.

I am storing the light state, oncoming traffic, traffic to the left, and the direction of the next waypoint direction. The first three determine what are legal moves to make. The fourth is a hint as to where the car should go next. Traffic to the right doesn't matter as long as everyone is obeying the traffic laws.

What changes do you notice in the agent's behavior?

It is now tending to go towards the target more often than not, and seems to be obeying traffic rules (at least as often as drivers around here do…)

Report what changes you made to your basic implementation of Q-Learning to achieve the final version of the agent. How well does it perform?

I started with alpha, gamma, and explore probability all set to 0.5 I first decreased the explore probability to make it more likely to use the learned route (there are a small number of states and plenty of time for it to learn something, so exploration isn't as important). I increased alpha as the game is pretty deterministic with regard to the transitions, so it should be able to trust that it is learning the right thing. Gamma was reduced based on the observation that the next state was usually independent on the action the agent takes at the current light.

Does your agent get close to finding an optimal policy, i.e. reach the destination in the minimum possible time, and not incur any penalties?

The car does seem to be usually finding the correct policy, but seems to have issues at the edge of the world. This is likely because the world seems to be doughnut shaped so if it goes off the edge it ends up on the other side. Yet the waypoints seem to be calculated as if the world has fixed edges, so if it takes an experimental turn off the edge, even though it is only one square further away it thinks it has to go all across the world.