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

Answer：It just move randomly like a headless chicken, sometimes it may luckily reach the target location but need many moves.

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

Answer：I pick 'next_waypoint' and 'light' as my state. 'next_waypoint' give a baisc direction to the target location and 'light' tell the agent whether take an action without punishment. I also try to take 'presence of cars' into consideration because it could give infomation to reduce punishment when take an action. But as a result it doesn't make a better result and take more time to exploration.

3、What changes do you notice in the agent's behavior

Answer：At first the agent still move like a headless chicken when it's not familiar with the environment. But quickly as the q table grows, it take less and less time to reach the target location

4、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?

Answer：I have tried many choice of learning rate, discount factor and exploration probability. I think the choice in my code is better compared to other choice. It reach the destination quickly when familiar with the environment. Sometimes it incur penaltied because as I mentioned before, I ignore 'presence of cars' in my state. But I'm satisfied with the current performance, maybe I'm crazy haha!