**QUESTION:** Observe what you see with the agent's behavior as it takes random actions. Does the***smartcab***eventually make it to the destination? Are there any other interesting observations to note?

The smartcab reaches the destination but it takes a long time and it is not the optimal path. We can also see that the car several times loses points in the top left information in the game window.

***QUESTION:*** *What states have you identified that are appropriate for modeling the****smartcab****and environment? Why do you believe each of these states to be appropriate for this problem?*

I considered that the “self.next\_waypoint” and the “inputs” should be used. I need to know my next step, the light, if there is a car coming in, from left and from right. The “deadline” is useless since it will not help arriving faster in the target position, so I didn’t declare it.

***OPTIONAL:*** *How many states in total exist for the****smartcab****in this environment? Does this number seem reasonable given that the goal of Q-Learning is to learn and make informed decisions about each state? Why or why not?*

The total states are (3x2x2x2x2=) 48. For the Q function it is dependent of the states and actions, there is 4 possible actions, so 192 possibilities. The number is reasonable to make the decisions, considering that is expected to train the algorithm it is possible to let it running and learn about the decisions to make.

***QUESTION:*** *What changes do you notice in the agent's behavior when compared to the basic driving agent when random actions were always taken? Why is this behavior occurring?*

At first the car was moving to the expected direction but he had to do a movement different than normal (turn or a turn that it has not done before) it would just stop. After I changed the q max formula from “greater than” to “greater or equal to” and it started working perfectly. It was happening because to be stopped would not generate and penalty and it would not update its reward. When I changed to “greater or equal to”, it would prefer to test anything than to be stopped (in the list possible\_actions, None is the first), unless it has already learned that anything else is bad.

**Improve the Q-Learning Driving Agent**

Your final task for this project is to enhance your driving agent so that, after sufficient training, the **smartcab** is able to reach the destination within the allotted time safely and efficiently. Parameters in the Q-Learning algorithm, such as the learning rate (alpha), the discount factor (gamma) and the exploration rate (epsilon) all contribute to the driving agent’s ability to learn the best action for each state. To improve on the success of your **smartcab**:

* Set the number of trials, n\_trials, in the simulation to 100.
* Run the simulation with the deadline enforcement enforce\_deadlineset to True (you will need to reduce the update delay update\_delayand set the display to False).
* Observe the driving agent’s learning and **smartcab’s** success rate, particularly during the later trials.
* Adjust one or several of the above parameters and iterate this process.

This task is complete once you have arrived at what you determine is the best combination of parameters required for your driving agent to learn successfully.

***QUESTION:*** *Report the different values for the parameters tuned in your basic implementation of Q-Learning. For which set of parameters does the agent perform best? How well does the final driving agent perform?*

***QUESTION:*** *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? How would you describe an optimal policy for this problem?*