**Udacity Smartcab project**

Tomoaki Tsuzuki

**Implement a Basic Driving Agent**

***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?*

***ANSWER:*** *I have run 100 times with random actions. The agent rarely reached the goal within the time with* ***enforce\_deadline*** *to be* ***True****. It is 16/100 times. The average reward for those trials is 0.2.*

*The graph below shows how random smartcab performed.*

*Blue line shows if the agent reached to destination (0 = do not reached to goal, 1 = reached to goal). Orange line shows net reward of the trial.*

*Obviously there is no trend for both success or fail and reward.*

*Hopefully with Q-learning implemented, it shows trend of getting better and better.*

### Inform the Driving Agent

***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?*

***Answer:*** *I have selected below as the state variables.*

* *light {red, green}*
* *oncoming {None, forward, left, right}*
* *left {None, forward, left, right}*
* *right {None, forward, left, right}*
* *next\_waypoint {None, forward, legt, right)*

*The state is combination of those state variables. So it is total of 512 states.*

*Since possible actions to take is {None, forward, left, right}, my Q-matrix would be the size of 512 x 4.*

*There is one more state variables that can be used, which is dead line.*

*However, it greatly increases the size of Q-matrix because deadline variable can take many states (distance between start and destination times 5).*

*For the agent to learn within the reasonable amount of time, I did not include deadline as state variables.*

### Implement a Q-Learning Driving Agent

***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?*

***Answer:*** *With the Q learning implemented, the agent starts to reach goal after few times of trial. This is expected as the agent has learnt how to take action for the first few trials and then trained well enough to get to the goal within the time limit.*

*The graph below shows how Q learning smartcab performed. Again, Blue line shows if the agent reached to destination and orange line shows net reward of trial.*

*There are great improvements seen by using Q learning. The agent gets to the goal 86/100 times and average reward is about 20. However, the agent does not get to the goal from time to time, and there is some room for improvement.*