**TRAINING A SMARTCAB**

***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 does eventually make it to the destination, although the trial run does occasionally hit the hard time limit. Ofcourse, there is a large variation in the number of steps the car takes to get to the target and the path is nowhere near optimal.
* The edges of the grid are a little different from the rest of the intersections as the car appears on the other side of the grid after crossing beyond the end (i.e. will appear on the left side if going through the wall on the right). This may cause some problems in training as the optimal path may involve using this to the cars advantage, but the intermediate rewards may not reflect this as it may register as moving away from the target in the short run.

***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? 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?***

* Different factors to consider:
  + 48 different intersections
  + Two types of lights, red and green.
  + Four possible headings of the car, North, South, East, West.
  + Oncoming traffic going left, right, forward or not there.
  + North-South difference to target: 5 possibilities
  + East-West difference to target: 7 possibilities
  + Time Remaining

This leads to a total of 53760 unique states at any given moment, ignoring time remaining, which would increase the possible states even more. However, not all these states need to be considered when training the smartcab.

* Relevant factors to consider:
  + Two types of lights, red and green
  + Four possible directions of car
  + Oncoming traffic going left, right, forward or not there.

Using just these possible factors, we get a total of 32 possible states at any given intersection. Using these factors, the smartcab should be able to use the rewards based on each selection to learn the optimal way to drive.