Smart Cab Project

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**: As the agent takes random behaviors, the agent eventually reaches the destination, but the time it takes to get there is very random.

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**: Some states I have identified as the location of the destination vs. the smartcab. I believe these states are appropriate for this problem because at farther distances, the smartcab can take more conservative actions where as the smartcab can focus on travelling the shortest distance if the location were closer.

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?

> **ANSWER**: There are 6*8*2 = 96 states. Since we are training the smartcab over 100 trials, we don't want too many states such that the model will never train over, nor too few such that it doesn't capture the complexity of the environment.

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:** Some changes are that the smartcab is learning to not take actions that resulted in a negative reward and taking actions that resulted in positive rewards when reaching an old state.

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?

ANSWER: I tried gamma, alpha, and epsilon values in the range of [.1, .7], [.1, .7], [.0, .05] respectively. What seemed to work best (consistently above .93%) was a gamma, alpha, and epsilon combination of 0.7, 0.1, and 0.0.

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?

ANSWER: Yes, I would say that my smartcab reaches the optimal policy. I would describe the optimal policy for the problem as the highest success rate with the least amount of time used. Over 100 trials, I have an average success rate of .93 and an average percent time used of around .08.