Project Report

You will be required to submit a project report along with your modified agent code as part of your submission. As you complete the tasks below, include thorough, detailed answers to each question *provided in italics*.

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?*

The smartcab performed quite poorly given random actions. It did eventually reach its destination on some trials, but well outside of the potential deadline constraint. Additionally, the smartcab's behavior seems to ignore traffic light states when only given random actions as inputs and no other information about the environment.

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?*

The states I identified as necessary for modeling the smartcab and the environment are:

1. The valid directions for the **next waypoint** from the planner [‘left’, ‘forward’, ‘right’] value is always the relative direction of the destination
2. The current status of the **traffic lights** [‘red’, ‘green’]
3. **Oncoming** informs the agent about the intended direction of the car on the oncoming lane will go next [‘None’, ‘forward’, ‘left’, right’]
4. Traffic to the **left** of the agent at the current intersection inputs[‘left’]

These three states are important because they are the key factors in determining the behavior of the smartcab and possible actions it could take based on the current environment at each intersection. For the **next waypoint** variable, it is necessary to include because much like how a GPS operates it is telling the agent the best action to take given its location relative to the destination. For the **traffic lights** variable, it is necessary to include to ensure the agent doesn't not violate any traffic laws by taking illegal actions while the light is red. The **oncoming** variable is necessary to include to ensure the agent obeys traffic laws when there is oncoming traffic present such as yielding to through traffic while making a left turn. Lastly, the presence of traffic from the **left** is necessary for the agent to know because if the agent wants to turn left it does not have the right of way and must wait for perpendicular traffic to pass through. Deadline could also be used as a state to model the environment but I chose not to include it to reduce the complexity of the state space. We can ignore traffic from the right directions because of traffic laws.

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?*

Once I implemented a Q-Learning driving agent I noticed the smartcab's was able to reach the destination within the time constraint of the deadline for each trial. When the smartcab encountered a red light it would sometimes circle around back to the destination instead of waiting for the light to cycle. Other times the smartcab would seem to take a non-optimal route an appear stuck due to local minima. Initial trials had a higher occurrence of failure under the deadline constraint as the smartcab takes some time initially to learn the environment before consistently reaching the destination successfully. As the agent learns it more often follows the best next waypoint and is making less random exploration actions.

Improve the Q-Learning Driving Agent

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?*

In order to test the performance of the Q-Learning algorithm; I decided to look at 3 metrics:

1. Success Rate (%)
   Number of times the Primary agent reached its destination over 100 trials.
2. Average Penalties Incurred
   Average number of times the Primary agent received a negative reward over 100 trials.
3. Efficiency (%)
   The percentage of deadline used in order to reach the destination. Lower is better.

For my initial parameter tuning I set gamma to 0 and incremented alpha.

alpha	gamma	Success Rate (%)	Avg. Penalty	Avg. Efficiency (%)
0.2	0.0	98%	0.43	45%
0.4	0.0	98%	0.26	47%
0.6	0.0	98%	0.19	46%
0.8	**0.0**	**99%**	**0.16**	**46%**
1	0.0	100%	0.23	45%

From that initial set of data, I took the alpha/gamma values I thought performed **best** given my metrics and then incremented gamma with a constant alpha value.

0.8	**0.2**	**100%**	**0.23**	**41%**
0.8	0.4	99%	0.31	45%
0.8	0.6	99%	0.39	42%
0.8	0.8	99%	0.35	44%
0.8	1.0	82%	1.47	64%

For my second iteration of parameter tuning I kept gamma constant and ran the trial twice with a slightly larger than a smaller alpha value. For the third iteration of parameter tuning I took the best performing alpha value and ran the trial twice more with a slightly larger than a slightly smaller gamma value with a constant alpha.

0.79	0.2	99%	0.21	45%
0.81	0.2	99%	0.20	43%
0.81	**0.19**	**100%**	**0.22**	**41%**
0.81	0.21	98%	0.22	48%

After several trials and parameter tuning iterations, according to my performance metrics an alpha = 0.81 and gamma = 0.19 come closest to an optimal policy with a 100% success rate, average penalties per trial less than 1 an efficiency of nearly 40%.

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?*

After several trials and parameter tuning iterations, according to my performance metrics an alpha = 0.81 and gamma = 0.19 come closest to an optimal policy with a 100% success rate, average penalties per trial less than 1 an efficiency of nearly %40. I think an optimal policy is one where the smartcab makes it to its destination within the alotted time or faster, does not commit any traffic violations and is able to do multiple trips with a high success rate. All factors which I think contribute greatly to the accumulation of an optimal smartcab driving policy.