

Qs: Student justifies a set of features that best model each state of the driving agent in the environment. Unnecessary features not included in the state (if applicable) are similarly justified.

You choose not to use deadline. Please provide a justification for doing so (a good idea) in this section.

Ans:

`deadline` is not included in the state space so we can have a practical size of the state space. Also, driving safely is the top priority of the Smartcab, and so it should not be guided based on how much time is left. This explanation has been added to the answer to Question 5.

Qs: You have some changes to make to your Q-learning algorithm. You actually made it *too* complex. A good thing in the long run, but should be modified for this project.

Ans:

The gamma term in Q-learning was removed (function `learn()` in `agent.py`).

Qs: Approximately how many training trials did the driving agent require before testing? Does that number make sense given the epsilon-tolerance?

Specifically, you mention 30 trials but the epsilon value will be 0 after 20.

Ans:

When the tolerance `epsilon` becomes zero, the driving agent always choose an action which gives a largest Q-score, that means the driving agent becomes confident in the knowledge it has been learnt and tends to apply that knowledge, or “exploiting”, instead of keep “exploring” (function “choose_action()” in `agent.py`). Although it has stopped “exploring” randomly, the Smartcab keeps learning by updating the Q-table (alpha is not zero). I think this setting makes sense.

Qs: A few things: Please address what it means to stop exploring after 20 trials but to continue learning. Why did you select an initial value of 20.0 for a new state? Why did you choose the epsilon-tolerance and alpha values that you did?

Ans:

The first question has been answered in the question above. The initial value of 20.0 for the new state and the choice of epsilon and alpha are based on experiments. I tried many different values for these parameters, and it turns out that the value here gives me a good result in terms of safety and reliability.