Smartcab Report

Alex Leu

In your report, mention what you see in the agent's behavior. Does it eventually make it to the target location?

The agent has the motion of either going none, forward, left and right. If there is an oncoming car, or the traffic light is red, it would prevent the agent from moving forward. The agent eventually makes it to the target locaiton when the dealine criteria is turned to "False."

Justify why you picked these set of states, and how they model the agent and its environment.

The state shows whether there are oncoming car, whether there is a green/red lights, and whether the oncoming is going left or right, and the next action of where the agent is going to. It can either go "none", "forward", "left", "right".

The directions were picked as they were the only valid actions the agents can take. The states were shown to be what could get affect the motion of the agent, referenced from the dummy agent.

By setting up as the deadline limit, it provides limited number of tries the agent can do to get to the destination. If we do not have the deadline limit, the agent will eventually reach to the point, as there is a finite grid of the table.

What changes do you notice in the agent's behavior?

It stays stationary a lot at first. After it picks up, and gets close to the target, it often turns in a circle. We can see the agent is learning, and trying to reach to the destination.

Report what changes you made to your basic implementation of Q-Learning to achieve the final version of the agent. How well does it perform? 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?

Here are the list of changes regarding to the constants for alpha, gamma, and epison. We went through the trial and erorr to find out the best tuning for the algorithmn. By changing/tuning over the values, we found the the tuning is at the most optimal at.

Here's the chart for the tuning:

alpha	gamma	episilon	deadline
0.25	0.25	0.25	-41
0.25	0.25	0.5	-83
0.25	0.25	0.75	-19
0.25	0.25	1	5
0.25	0.5	0.25	-23
0.25	0.5	0.5	24
0.25	0.5	0.75	-6
0.25	0.5	1	-45
0.25	0.75	0.25	-5
0.25	0.75	0.5	-14
0.25	0.75	0.75	-183
0.25	0.75	1	-6
0.25	1.0	0.25	20
0.5	0.5	0.5	-44
0.75	0.5	0.5	-39
0.75	0.75	0.5	8
0.75	0.75	0.75	-3
1	0.75	0.75	-6
1	1	0.75	-112
1	1	1	-82

In [ ]: