REPORT

TRAIN A SMARTCAB TO DRIVE

USING Q-LEARNING TO TRAIN A SMART CAB TO DRIVE AROUND IDEALISTIC CITY ROADS

ADITYA HARSH

7/20/2016

REPORT

TRAIN A SMARTCAB TO DRIVE

1) Implement a Basic Driving Agent

- **What you see in the agent's behavior. Does it eventually make it to the target location?**
  - In order to implement a basic driving agent, I utilized random.choice over a list of the possible actions. The agent does eventually make it to the target location, but it's just due to random chance.

2) Identify and Update State

- **Identify a set of states that are appropriate for modeling the driving state.**
  - To model my driving agent, I chose to use the color of the light, the presence of an oncoming car, whether or not there is a car present to the left and right of our car, and the next waypoint.

- **Why you picked this set of states and how they model the agent and its environment?**
  - My agent is concerned with making it to the goal state as quickly as possible, while avoiding collisions and other negative consequences that are built into the rewards system. Because the path planning is already implemented for us, this means that we simply want our agent to decide at each step whether it should attempt to move directly toward the next waypoint, or if there are obstacles that prevent this from being a feasible strategy. In order to do that, our state must capture both the waypoint and any potential obstacles, so those are the features that I chose to include in my state space.

3) Implement Q-Learning

- **What changes did you notice in the Agent's behavior?**
  - After implementing Q-Learning to inform the agent's policy behavior. The agent mills around for a while it is in the initial stages of the learning process, but quickly begins to direct its movement towards the goal state.

4) Improve the Q-Learning Driving Agent

- **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?

- To model my driving agent, I chose to use the color of the light, the presence of an oncoming car, whether or not there is a car present to the left and right of our car, and the next waypoint.

- **Why you picked this set of states and how they model the agent and its environment?**
  - Tuning parameter includes:

    (1) Learning rate α: in our case, c and b are the tuning learning rate α. When α is close to 0, the agent tends to learn little from the interaction. When α is close to 1, the agent tends to have a "short memory". It reflects more on the recent interactions.

    (2) Discount factor γ: the value ranges from 0 to 1. The value more close to 0 the influence of future states are reduced and vice versa.

By Trial and error, we find by decreasing γ the agent has better change getting to the destination. The ideal γ is about 0.3 through experiments.