

Machine Learning NanoDegree

TRAIN A SMARTCAB TO DRIVE

IMPLEMENT A BASIC DRIVING AGENT

by OMOJU MILLER



August 10, 2016

DEFINITION

PROBLEM STATEMENT

In the not-so-distant future, taxicab companies across the United States no longer employ human drivers to operate their fleet of vehicles. Instead, the taxicabs are operated by self-driving agents—known as **smartcabs**—to transport people from one location to another within the cities those companies operate. In major metropolitan areas, such as Chicago, New York City, and San Francisco, an increasing number of people have come to rely on **smartcabs** to get to where they need to go as safely and efficiently as possible. Although **smartcabs** have become the transport of choice, concerns have arisen that a self-driving agent might not be as safe or efficient as human drivers, particularly when considering city traffic lights and other vehicles. To alleviate these concerns, your task as an employee for a national taxicab company is to use reinforcement learning techniques to construct a demonstration of a smartcab operating in real-time to prove that both safety and efficiency can be achieved.

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?

At first, I set the action of the self-driving agent as forward so I could understand how the grid world worked.

```
# Select action according to your policy  
action = random.choice(['forward'])
```

From this choice, it was evident the smartcab would never get to the destination, except in the case when we were lucky, and the goal just happened to be in an intersection ahead of the cab.

```
# Select action according to your policy  
action = random.choice([None, 'forward', 'left', 'right'])
```

It was frustrating to watch. Most of the time, the agent never made it to the destination. There were a few times when it did; however, as expected, it *didn't learn* anything from making good decisions or bad decisions. There were no consequences for its behavior.

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?

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?

MARKOV DECISION PROCESSES

States and transitions

The rewards are a means of capturing domain knowledge.

SEQUENCE OF REWARDS

- Infinite horizons
- Utility of sequences
- Additive assumption:
if
 $U(S_0, S_1, S_2, \dots) > U(S_0, S'_1, S'_2, \dots)$
then
 $U(S_1, S_2, \dots) > U(S'_1, S'_2, \dots)$

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?

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?
- 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?

CONCLUSION
