Smartcab

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

Random agent behavior has been implemented by following code:

```
# TODO: Update state
action_okay = True
if self.next_waypoint == 'right':
  if inputs['light'] == 'red' and inputs['left'] == 'forward':
    action_okay = False
elif self.next_waypoint == 'forward':
  if inputs['light'] == 'red':
    action_okay = False
elif self.next_waypoint == 'left':
  if inputs['light'] == 'red' or (inputs['oncoming'] == 'forward' or inputs['oncoming'] == 'right'):
    action_okay = False
action = None
if action_okay:
  action = self.next_waypoint
  self.next_waypoint = random.choice(Environment.valid_actions[1:])
```

Agent initialized and moved to the target location.

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

As a set of states I've decided to choose coordinates of agent (x,y). It will allow me later define Q function which will assign utility value with each state. Based on this Q function

it will be possible to define policy for each state. The state allow to identify position of agent in the environment relative to target.

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

On different steps of Q learning algorithm implementation I've observed different changes:

1) I've implement algorithm without gamma factor in Q formula. The car was cycled by boxes.
2) When I introduced gamma factor the car starts to reach destination without stopping (as in default planner).

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?

I make the following adjustments:
1) Make penalties for actions out of borders or not in line with traffic rules.

2) Penalties for staying on the place without moving

3) Add Q +1 for planner recommendation

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?

Yes: instead of staying on the crossroad (as it was i the default planner) my agent is moving based on Q values.