FAI Project Milestone 2

Project Name: Self-driving Smart Cab

Proposed Milestone Deliverables:

- Finish Q-Learning algorithm and have a working agent smart cab. Also, we will be looking forward to extending the project if feasible.

Our Progress:

- We have implemented a basic Q-learning agent. Most of the times it is taking the right action. However, we have noticed a few cases where it is not taking the correct action. For example, in some cases the agent is not following the waypoint. It is not breaking any traffic rules, but it is deviating from the waypoint path. In three cases, the agent also broke a traffic light with a big negative reward. So, we are looking to update our agent significantly to avoid these occurrences. Below we have posted the

  Simulating trial. . .
  Agent not set to learn.

  Step 0 Results:

  Agent previous state: ['forward', 'forward', None, None]
  Agent drove right instead of forward. (rewarded 1.93)
  Agent not enforced to meet deadline.

  Step 1 Results:

  Agent previous state: ['left', 'forward', 'forward', None]
  Agent drove right instead of left. (rewarded 1.84)
  Agent not enforced to meet deadline.

  Step 2 Results:

  Agent previous state: ['forward', None, None, 'forward']
  Agent properly idled at a red light. (rewarded 2.93)
  Agent not enforced to meet deadline.

Step 3 Results:

Agent previous state: ['forward', None, None, 'forward']
Agent properly idled at a red light. (rewarded 1.50)
Agent not enforced to meet deadline.

Step 4 Results:

Agent previous state: ['forward', 'right', None, 'forward']
Agent drove right instead of forward. (rewarded 1.16)
Agent not enforced to meet deadline.

Step 5 Results:

Agent previous state: ['left', None, None, None]
Agent drove right instead of left. (rewarded 1.21)
Agent not enforced to meet deadline.

Step 6 Results:

Agent previous state: ['forward', None, 'left', None]
Agent idled at a green light with no oncoming traffic. (rewarded -4.90)
Agent not enforced to meet deadline.

Step 7 Results:

Agent previous state: ['forward', 'forward', 'right', 'left']
Agent followed the waypoint forward. (rewarded 1.42)
Agent not enforced to meet deadline.

Step 8 Results:

Agent previous state: ['forward', None, 'forward', 'right']
Agent attempted driving left through traffic and cause a minor accident. (rewarded -20.60)
Agent not enforced to meet deadline.

Step 9 Results:

Agent previous state: ['forward', None, 'forward', None]
Agent followed the waypoint forward. (rewarded 1.07)
Agent not enforced to meet deadline.

Step 10 Results:

Agent previous state: ['forward', 'forward', None, 'forward']
Agent followed the waypoint forward. (rewarded 2.63)
Agent not enforced to meet deadline.

Step 11 Results:

Agent previous state: ['forward', None, None, 'left']
Agent drove left instead of forward. (rewarded 0.12)
Agent not enforced to meet deadline.

Step 12 Results:

Agent previous state: ['right', None, 'right', None]
Agent attempted driving forward through a red light. (rewarded -9.53)
Agent not enforced to meet deadline.

Step 13 Results:

Agent previous state: ['right', None, None, None]
Agent idled at a green light with no oncoming traffic. (rewarded -4.17)
Agent not enforced to meet deadline.

Step 14 Results:

Agent previous state: ['right', None, None, 'left']
Agent drove left instead of right. (rewarded 0.42)
Agent not enforced to meet deadline.

Step 15 Results:

Agent previous state: ['right', None, None, 'right']
Agent attempted driving left through a red light with traffic and cause a major accident. (rewarded -39.20)

Agent not enforced to meet deadline.

Step 16 Results:

Agent previous state: ['right', None, None, 'right']
Agent drove forward instead of right. (rewarded 0.80)
Agent not enforced to meet deadline.

Step 17 Results:

Agent previous state: ['right', 'forward', 'right', None]
Agent followed the waypoint right. (rewarded 2.19)
Agent not enforced to meet deadline.

Step 18 Results:

Agent previous state: ['right', None, None, 'forward']
Agent followed the waypoint right. (rewarded 2.10)
Agent not enforced to meet deadline.

Step 19 Results:

Agent previous state: ['forward', None, None, None]
Agent drove left instead of forward. (rewarded 0.46)
Agent not enforced to meet deadline.

Step 20 Results:

Agent previous state: ['right', 'left', None, None]
Agent drove left instead of right. (rewarded 0.65)
Agent not enforced to meet deadline.

Step 21 Results:

Agent previous state: ['right', None, 'left', 'right']
Agent properly idled at a red light. (rewarded 1.30)
Agent not enforced to meet deadline.

Step 22 Results:

Agent previous state: ['right', None, 'forward', None]
Agent followed the waypoint right. (rewarded 1.64)
Agent not enforced to meet deadline.

Simulation ended. . .

So, as we can see here, there were three instances where the agent caused traffic infractions or was part of a collision. This is something that we have not yet been able to resolve. Not only that, at most of the steps the agent is not taking the optimal route which would make the rewards smaller. For the final project it is our goal that our agent can run through the simulation taking the best action at every step.

Milestone 2 Shortcomings:

We could not have an optimal Q-learning agent working as per our goals in this milestone. We are looking into this and will be improving upon it to the best of our ability before the final project presentation. For the final presentation, we also want to implement a graphical representation of our results using matplotlib. We are currently working on this too.