다음 코드는
Taxi-v3 게임을 실행하고,
10번 임의의 action을 수행해 보는 프로그램입니다.

```
import gym
env = gym.make("Taxi-v3")

print (env.observation_space.n ) #총 상태 수
print (env.action_space.n )      #action 수

for _ in range(1):  # 게임 횟수 1

    observation = env.reset()
    env.render()
    for _ in range(10):  #10 번 임의의 action 을 수행
            action = env.action_space.sample() # your agent here
            #아무 action 이나 수행하는 함수. (this takes random actions)
            observation, reward, done, info = env.step(action)
            print( "action",  action, "reward", reward)
            env.render()
```

Taxi-v3 게임 에 대한 다음 설명을 읽고,
Taxi-v3 게임을 자동을 실행하는 Agent를 만들어보세요. Q 훈련 후 게임의 성공확률을 높여보세요.

```
    The Taxi Problem
    from "Hierarchical Reinforcement Learning with the MAXQ Value Function Decomposition"
    by Tom Dietterich


Description:
    There are four designated locations in the grid world indicated by R(ed), B(lue), G(reen),
and Y(ellow). When the episode starts, the taxi starts off at a random square and the
passenger is at a random location. The taxi drive to the passenger's location, pick up the
passenger, drive to the passenger's destination (another one of the four specified locations),
and then drop off the passenger. Once the passenger is dropped off, the episode ends.


Observations:
    There are 500 discrete states since there are 25 taxi positions, 5 possible locations of
the passenger (including the case when the passenger is the taxi), and 4 destination
locations.

    Actions:
    There are 6 discrete deterministic actions:
    - 0: move south
    - 1: move north
    - 2: move east
    - 3: move west
    - 4: pickup passenger
```

```
    - 5: dropoff passenger

    Rewards:
    There is a reward of -1 for each action and an additional reward of +20 for delievering the
passenger. There is a reward of -10 for executing actions "pickup" and "dropoff" illegally.


    Rendering:
    - blue: passenger
    - magenta  : destination
    - yellow: empty taxi
    - green: full taxi
    - other letters: locations
```

your job is to pick up the passenger at one location and drop him off in another. You receive +20 points for a successful dropoff

`env.observation_space.n` 총 상태수 : 500

`env.action_space.n` action 갯수 : 6

한번 이동할 때마다 reward -1 :
Drop off에 실패하면 -10
Pickup 에 실패하면 -10