

CS6700: Reinforcement Learning

Tutorial 4: SARSA and Q-Learning

Gautham Govind A, EE19B022

```
In [1]: import numpy as np
import matplotlib.pyplot as plt
from tqdm import tqdm
from IPython.display import clear_output
%matplotlib inline
```

Problem Statement

In this section we will implement tabular SARSA and Q-learning algorithms for a grid world navigation task.

Environment details

The agent can move from one grid coordinate to one of its adjacent grids using one of the four actions: UP, DOWN, LEFT and RIGHT. The goal is to go from a randomly assigned starting position to goal position.

Actions that can result in taking the agent off the grid will not yield any effect. Lets look at the environment.

```
In [2]: DOWN = 0
UP = 1
LEFT = 2
RIGHT = 3
actions = [DOWN, UP, LEFT, RIGHT]
```

Let us construct a grid in a text file.

```
In [3]: from google.colab import drive
drive.mount('/content/drive')
```

Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).

```
In [4]: %%shell
cd drive/MyDrive/CS6700/CS6700_Tutorial_4_QLearning_SARSA
cat grid_world2.txt
```

```
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0
1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0
1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0
1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0
1 1 1 1 1 1 2 2 2 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 2 2 2 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 2 2 2 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 2 2 2 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 2 2 2 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 2 2 2 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

Out[4]:

This is a 17×23 grid. The reward when an agent goes to a cell is negative of the value in that position in the text file (except if it is the goal cell). We will define the goal reward as 100. We will also fix the maximum episode length to 10000.

Now let's make it more difficult. We add stochasticity to the environment: with probability 0.2 agent takes a random action (which can be other than the chosen action). There is also a westerly wind blowing (to the right). Hence, after every time-step, with probability 0.5 the agent also moves an extra step to the right.

Now let's plot the grid world.

```
In [5]: world = 'grid_world2.txt'
goal_reward = 100
```

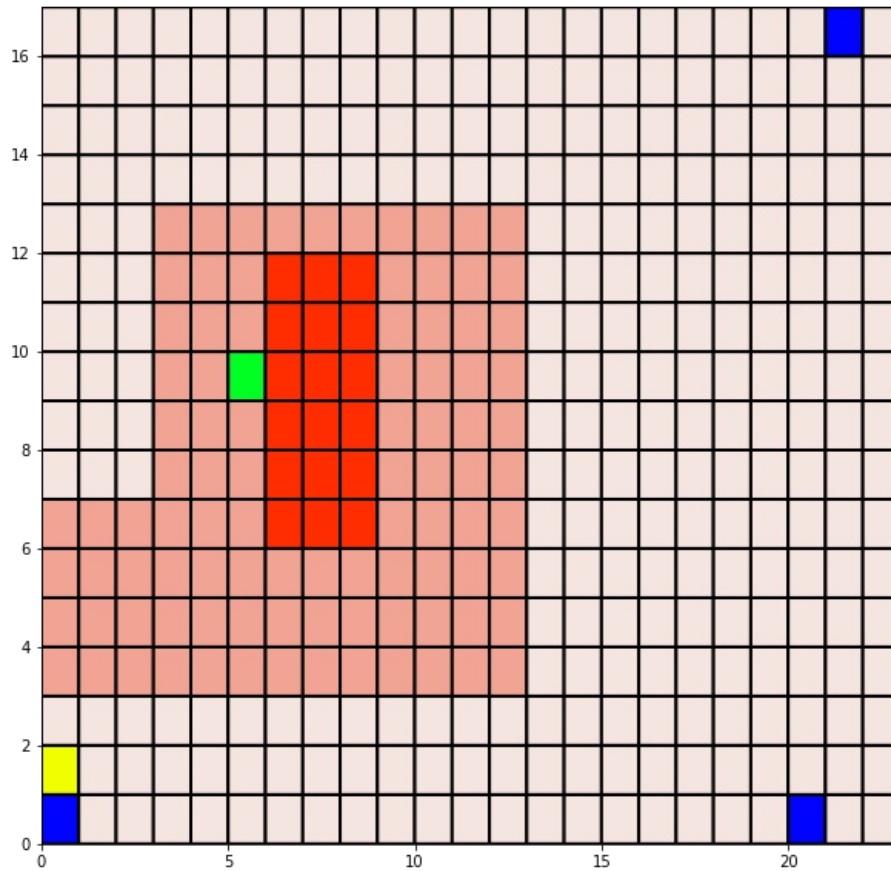
```

start_states = [(0,0), (0,20), (16,21)]
goal_states=[(9,5)]
max_steps=10000

from grid_world import GridWorldEnv, GridWorldWindyEnv

env = GridWorldEnv(world, goal_reward=goal_reward, start_states=start_states, goal_states=goal_states,
                    max_steps=max_steps, action_fail_prob=0.2)
plt.figure(figsize=(10, 10))
# Go UP
env.step(UP)
env.render(ax=plt, render_agent=True)

```



Legend

- *Blue* is the **start state**.
- *Green* is the **goal state**.
- *Yellow* is current **state of the agent**.
- *Redness* denotes the extent of **negative reward**.

Q values

We can use a 3D array to represent Q values. The first two indices are X, Y coordinates and last index is the action.

```

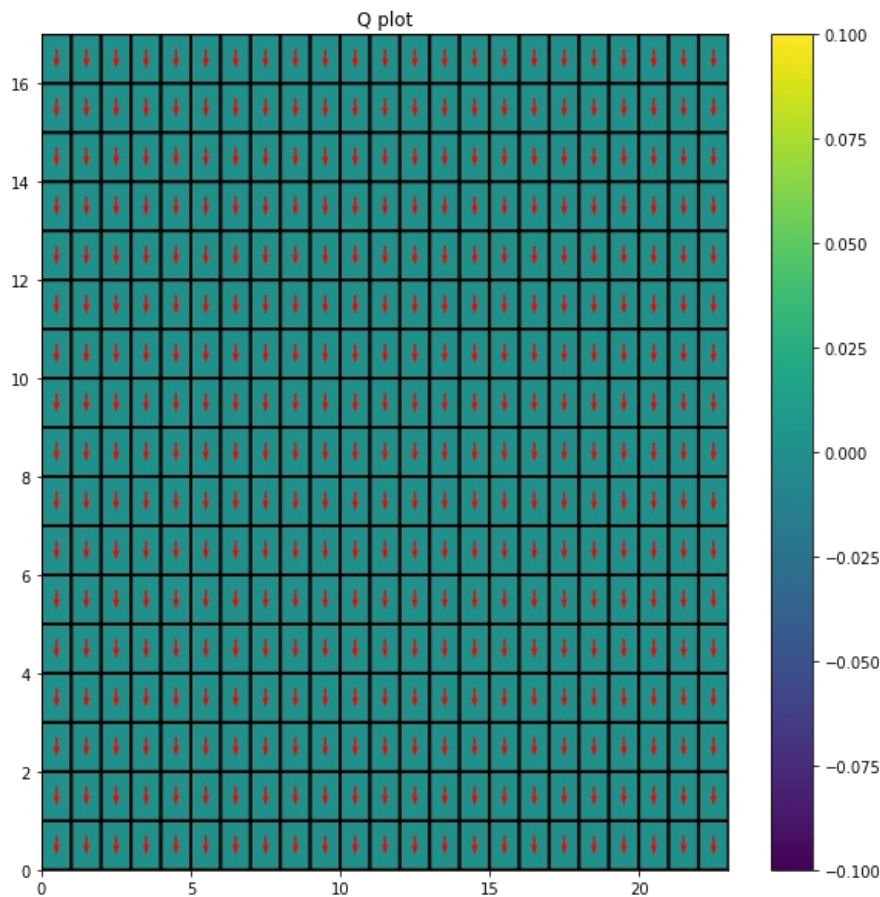
In [6]: from grid_world import plot_Q

Q = np.zeros((env.grid.shape[0], env.grid.shape[1], len(env.action_space)))

plot_Q(Q)

Q.shape

```



Out[6]: (17, 23, 4)

Exploration strategies

1. Epsilon-greedy
2. Softmax

```
In [7]: from scipy.special import softmax

seed = 42
rg = np.random.RandomState(seed)

# Epsilon greedy
def choose_action_epsilon(Q, state, epsilon, rg=rg):
    if ( np.random.rand() < epsilon ): # TODO: eps greedy condition
        return np.random.choice(np.array([0, 1, 2, 3]))
    else:
        return np.argmax(Q[state[0], state[1], :])# TODO: return best action

# Softmax
def choose_action_softmax(Q, state, rg=rg):
    return np.random.choice(np.array([0, 1, 2, 3]), p = softmax(Q[state[0], state[1], :]))
```

SARSA

Now we implement the SARSA algorithm.

Recall the update rule for SARSA:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_t + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)]$$

Hyperparameters

So we have some hyperparameters for the algorithm:

- α
- number of *episodes*.
- ϵ : For epsilon greedy exploration

```
In [8]: # initialize Q-value
Q = np.zeros((env.grid.shape[0], env.grid.shape[1], len(env.action_space)))
```

```
alpha0 = 0.4
gamma = 0.9
episodes = 10000
epsilon0 = 0.1
```

Let's implement SARSA

```
In [9]: print_freq = 100

def sarsa(env, Q, gamma = 0.9, plot_heat = False, choose_action = choose_action_softmax):

    episode_rewards = np.zeros(episodes)
    steps_to_completion = np.zeros(episodes)
    if plot_heat:
        clear_output(wait=True)
        plot_Q(Q)
    epsilon = epsilon0
    alpha = alpha0
    for ep in tqdm(range(episodes)):
        tot_reward, steps = 0, 0

        # Reset environment
        state = env.reset()
        action = choose_action(Q, state)
        done = False
        while not done:
            state_next, reward, done = env.step(action)
            action_next = choose_action(Q, state_next)

            Q[state[0], state[1], action] = Q[state[0], state[1], action] + alpha0*(reward + gamma*Q[state_next[0], state_next[1], action_next])

            tot_reward += reward
            steps += 1

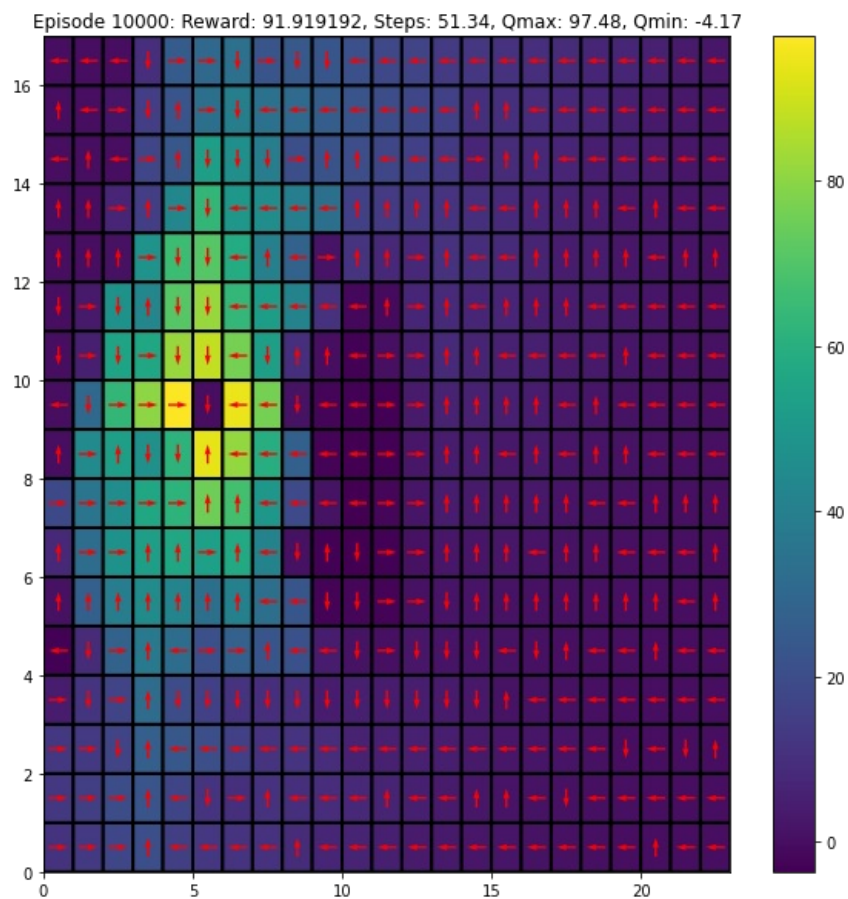
            state, action = state_next, action_next

        episode_rewards[ep] = tot_reward
        steps_to_completion[ep] = steps

        if (ep+1)%print_freq == 0 and plot_heat:
            clear_output(wait=True)
            plot_Q(Q, message = "Episode %d: Reward: %f, Steps: %.2f, Qmax: %.2f, Qmin: %.2f"%(ep+1, np.mean(episode_rewards[ep-print_freq:ep]), np.mean(steps_to_completion[ep-print_freq:ep]), Q.max(), Q.min()))

    return Q, episode_rewards, steps_to_completion
```

```
In [10]: Q, rewards, steps = sarsa(env, Q, gamma = gamma, plot_heat=True, choose_action= choose_action_softmax)
```



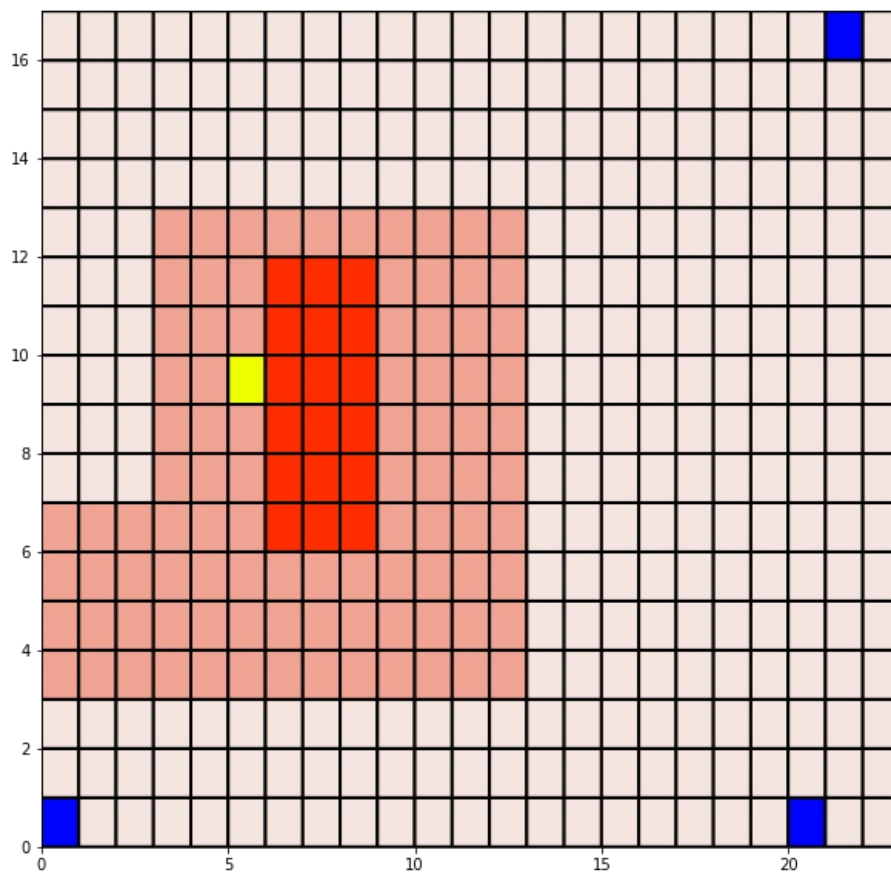
100% |██████████| 10000/10000 [02:01<00:00, 82.42it/s]

Visualizing the policy

Now let's see the agent in action. Run the below cell (as many times) to render the policy;

```
In [11]: from time import sleep

state = env.reset()
done = False
steps = 0
tot_reward = 0
while not done:
    clear_output(wait=True)
    state, reward, done = env.step(Q[state[0], state[1]].argmax())
    plt.figure(figsize=(10, 10))
    env.render(ax=plt, render_agent=True)
    plt.show()
    steps += 1
    tot_reward += reward
    sleep(0.2)
print("Steps: %d, Total Reward: %d"%(steps, tot_reward))
```



Steps: 23, Total Reward: 77

Analyzing performance of the policy

We use two metrics to analyze the policies:

1. Average steps to reach the goal
2. Total rewards from the episode

To ensure, we account for randomness in environment and algorithm (say when using epsilon-greedy exploration), we run the algorithm for multiple times and use the average of values over all runs.

```
In [12]: num_expts = 5
reward_avgs, steps_avgs = [], []

for i in range(num_expts):
    print("Experiment: %d"%(i+1))
    Q = np.zeros((env.grid.shape[0], env.grid.shape[1], len(env.action_space)))
    rg = np.random.RandomState(i)

    Q, rewards, steps = sarsa(env, Q, gamma = gamma, plot_heat=False, choose_action= choose_action_softmax)
    reward_avgs.append(rewards)
    steps_avgs.append(steps)

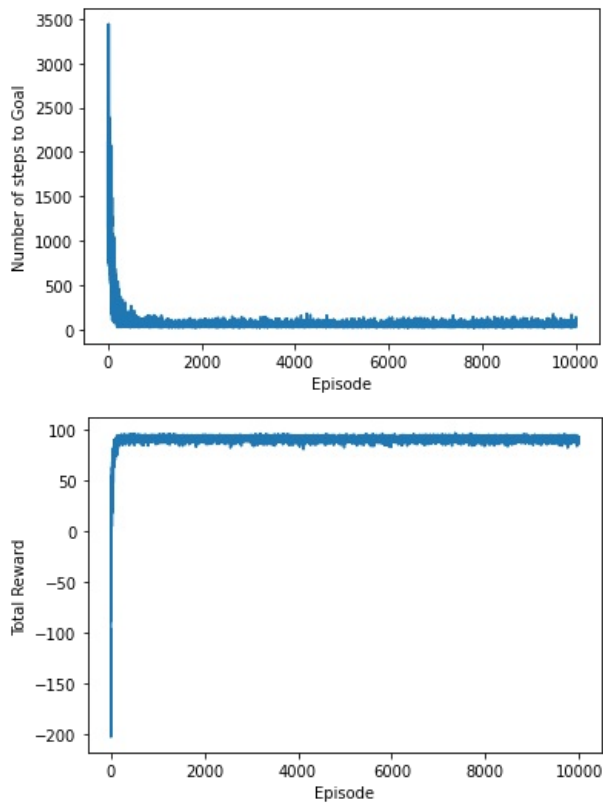
    # TODO: run sarsa, store metrics
```

```
Experiment: 1
100%|██████████| 10000/10000 [01:44<00:00, 95.78it/s]
Experiment: 2
100%|██████████| 10000/10000 [01:02<00:00, 159.73it/s]
Experiment: 3
100%|██████████| 10000/10000 [02:00<00:00, 83.24it/s]
Experiment: 4
100%|██████████| 10000/10000 [01:07<00:00, 149.09it/s]
Experiment: 5
100%|██████████| 10000/10000 [00:57<00:00, 174.80it/s]
```

```
In [13]: plt.figure()
plt.plot(list(range(10000)), np.mean(np.array(steps_avgs), axis = 0))
plt.xlabel('Episode')
plt.ylabel('Number of steps to Goal')
plt.show()

plt.figure()
plt.plot(list(range(10000)), np.mean(np.array(reward_avgs), axis = 0))
plt.xlabel('Episode')
```

```
plt.ylabel('Total Reward')
plt.show()
```



Q-Learning

Now, implement the Q-Learning algorithm as an exercise.

Recall the update rule for Q-Learning: $Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_t + \gamma \max_{a'} Q(s_t, a') - Q(s_t, a_t)]$

Visualize and compare results with SARSA.

```
In [14]: # initialize Q-value
Q = np.zeros((env.grid.shape[0], env.grid.shape[1], len(env.action_space)))

alpha0 = 0.4
gamma = 0.9
episodes = 10000
epsilon0 = 0.1
```

```
In [15]: print_freq = 100

def qlearning(env, Q, gamma = 0.9, plot_heat = False, choose_action = choose_action_softmax):

    episode_rewards = np.zeros(episodes)
    steps_to_completion = np.zeros(episodes)
    if plot_heat:
        clear_output(wait=True)
        plot_Q(Q)
    epsilon = epsilon0
    alpha = alpha0
    for ep in tqdm(range(episodes)):
        tot_reward, steps = 0, 0

        # Reset environment
        state = env.reset()
        action = choose_action(Q, state)
        done = False
        while not done:
            state_next, reward, done = env.step(action)
            action_next = choose_action(Q, state_next)

            # TODO: update equation

            Q[state[0], state[1], action] = Q[state[0], state[1], action] + alpha0*(reward + gamma*np.max(Q[state[0], state[1], :]) - Q[state[0], state[1], action])

            tot_reward += reward
            steps += 1

            state, action = state_next, action_next
```



```

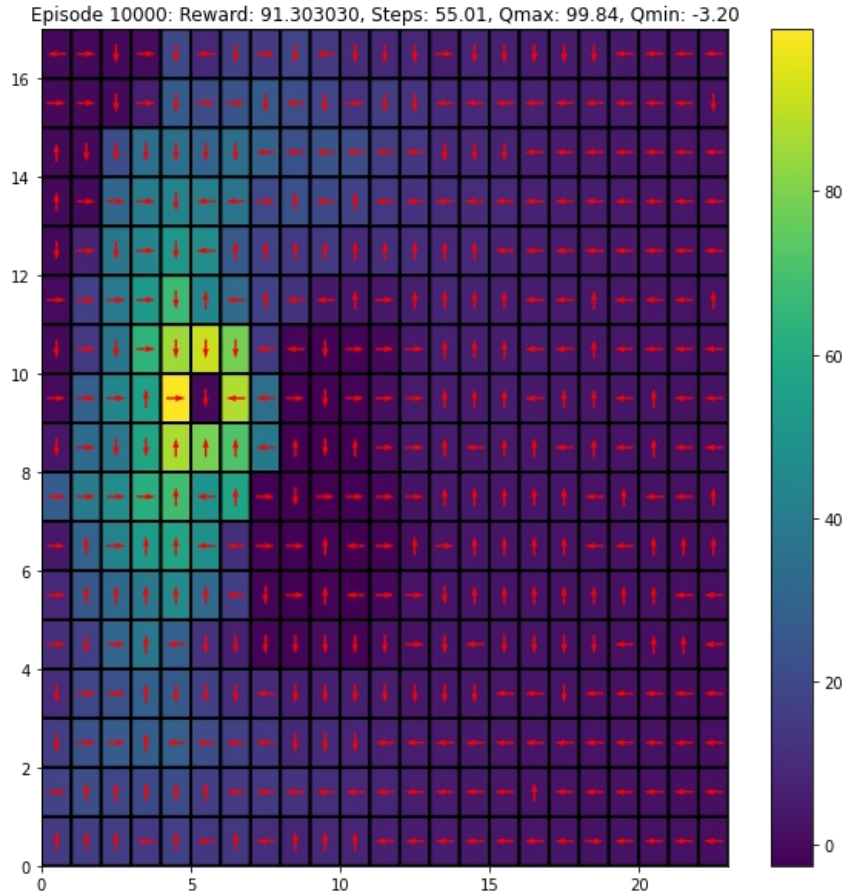
episode_rewards[ep] = tot_reward
steps_to_completion[ep] = steps

if (ep+1)%print_freq == 0 and plot_heat:
    clear_output(wait=True)
    plot_Q(Q, message = "Episode %d: Reward: %f, Steps: %.2f, Qmax: %.2f, Qmin: %.2f"%(ep+1, np.mean(ep.
np.mean(steps_to_completion[ep-print_
Q.max(), Q.min()))

return Q, episode_rewards, steps_to_completion

```

In [16]: `Q, rewards, steps = qlearning(env, Q, gamma = gamma, plot_heat=True, choose_action= choose_action_softmax)`



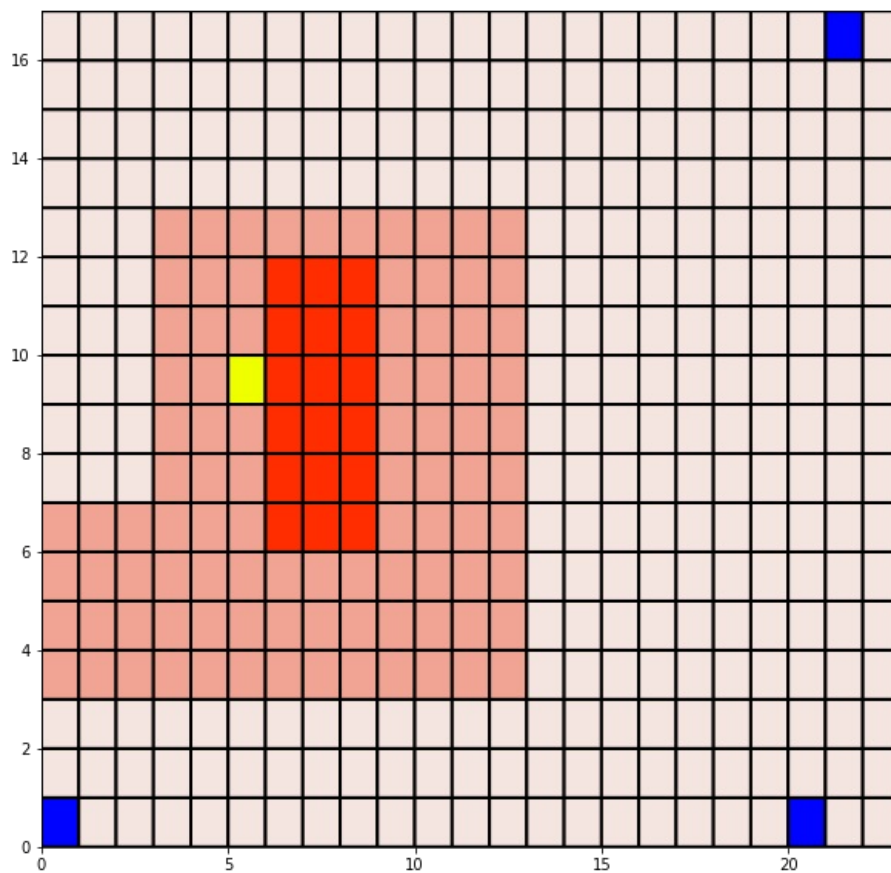
100%|██████████| 10000/10000 [02:04<00:00, 80.05it/s]

```

In [17]: from time import sleep

state = env.reset()
done = False
steps = 0
tot_reward = 0
while not done:
    clear_output(wait=True)
    state, reward, done = env.step(Q[state[0], state[1]].argmax())
    plt.figure(figsize=(10, 10))
    env.render(ax=plt, render_agent=True)
    plt.show()
    steps += 1
    tot_reward += reward
    sleep(0.2)
print("Steps: %d, Total Reward: %d"%(steps, tot_reward))

```

Steps: 36, Total Reward: 92

```
In [18]: num_expts = 5
reward_avgs, steps_avgs = [], []

for i in range(num_expts):
    print("Experiment: %d"%(i+1))
    Q = np.zeros((env.grid.shape[0], env.grid.shape[1], len(env.action_space)))
    rg = np.random.RandomState(i)

    Q, rewards, steps = qlearning(env, Q, gamma = gamma, plot_heat=False, choose_action= choose_action_softmax)
    reward_avgs.append(rewards)
    steps_avgs.append(steps)
    # TODO: run qlearning, store metrics
```

Experiment: 1

100%|██████████| 10000/10000 [01:14<00:00, 134.89it/s]

Experiment: 2

100%|██████████| 10000/10000 [01:07<00:00, 148.88it/s]

Experiment: 3

100%|██████████| 10000/10000 [01:14<00:00, 134.98it/s]

Experiment: 4

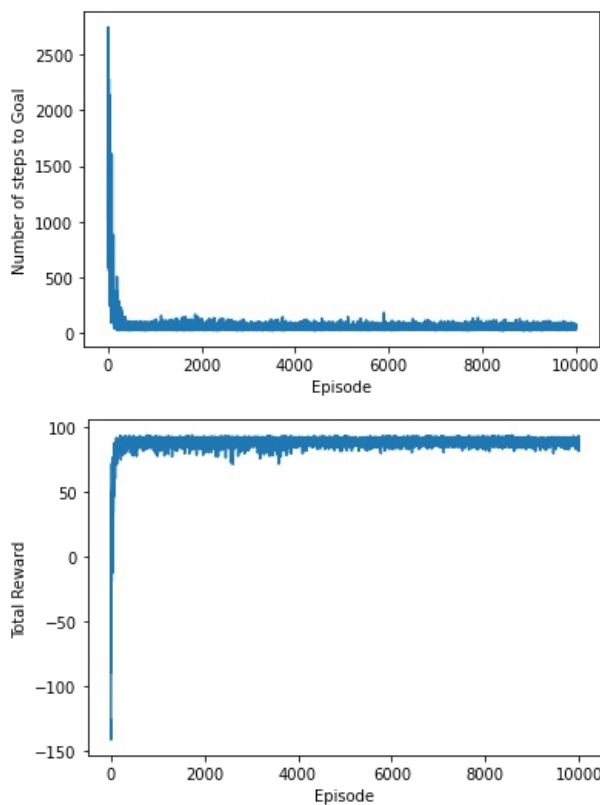
100%|██████████| 10000/10000 [00:58<00:00, 171.32it/s]

Experiment: 5

100%|██████████| 10000/10000 [01:53<00:00, 88.45it/s]

```
In [19]: plt.figure()
plt.plot(list(range(10000)), np.mean(np.array(steps_avgs), axis = 0))
plt.xlabel('Episode')
plt.ylabel('Number of steps to Goal')
plt.show()

plt.figure()
plt.plot(list(range(10000)), np.mean(np.array(reward_avgs), axis = 0))
plt.xlabel('Episode')
plt.ylabel('Total Reward')
plt.show()
```



TODO: What differences do you observe between the policies learnt by Q Learning and SARSA (if any).

From the Q-value plots, it can be observed that Q-Learning exhibits a more aggressive behaviour as compared to SARSA which tends to be more conservative. This basically means that SARSA tends to be more risk averse in the sense that the regions giving negative rewards tend to get avoided a lot (essentially because it is on-policy learning), whereas in Q-Learning, since we are taking max over all actions (off-policy learning), the policy is usually more aggressive.

```
In [20]: !pip install nbconvert
!sudo apt-get install texlive-xetex texlive-fonts-recommended texlive-plain-generic
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: nbconvert in /usr/local/lib/python3.8/dist-packages (5.6.1)
Requirement already satisfied: bleach in /usr/local/lib/python3.8/dist-packages (from nbconvert) (6.0.0)
Requirement already satisfied: Jinja2>=2.4 in /usr/local/lib/python3.8/dist-packages (from nbconvert) (2.11.3)
Requirement already satisfied: defusedxml in /usr/local/lib/python3.8/dist-packages (from nbconvert) (0.7.1)
Requirement already satisfied: testpath in /usr/local/lib/python3.8/dist-packages (from nbconvert) (0.6.0)
Requirement already satisfied: pandocfilters>=1.4.1 in /usr/local/lib/python3.8/dist-packages (from nbconvert) (1.5.0)
Requirement already satisfied: nbformat>=4.4 in /usr/local/lib/python3.8/dist-packages (from nbconvert) (5.7.3)
Requirement already satisfied: entrypoints>=0.2.2 in /usr/local/lib/python3.8/dist-packages (from nbconvert) (0.4)
Requirement already satisfied: jupyter-core in /usr/local/lib/python3.8/dist-packages (from nbconvert) (5.2.0)
Requirement already satisfied: traitlets>=4.2 in /usr/local/lib/python3.8/dist-packages (from nbconvert) (5.7.1)
Requirement already satisfied: pygments in /usr/local/lib/python3.8/dist-packages (from nbconvert) (2.6.1)
Requirement already satisfied: mistune<2, >=0.8.1 in /usr/local/lib/python3.8/dist-packages (from nbconvert) (0.8.4)
Requirement already satisfied: MarkupSafe>=0.23 in /usr/local/lib/python3.8/dist-packages (from Jinja2>=2.4->nbconvert) (2.0.1)
Requirement already satisfied: jsonschema>=2.6 in /usr/local/lib/python3.8/dist-packages (from nbformat>=4.4->nbconvert) (4.3.3)
Requirement already satisfied: fastjsonschema in /usr/local/lib/python3.8/dist-packages (from nbformat>=4.4->nbconvert) (2.16.2)
Requirement already satisfied: webencodings in /usr/local/lib/python3.8/dist-packages (from bleach->nbconvert) (0.5.1)
Requirement already satisfied: six>=1.9.0 in /usr/local/lib/python3.8/dist-packages (from bleach->nbconvert) (1.15.0)
Requirement already satisfied: platformdirs>=2.5 in /usr/local/lib/python3.8/dist-packages (from jupyter-core->nbconvert) (3.0.0)
Requirement already satisfied: attrs>=17.4.0 in /usr/local/lib/python3.8/dist-packages (from jsonschema>=2.6->nbformat>=4.4->nbconvert) (22.2.0)
Requirement already satisfied: importlib-resources>=1.4.0 in /usr/local/lib/python3.8/dist-packages (from jsonschema>=2.6->nbformat>=4.4->nbconvert) (5.10.2)
Requirement already satisfied: pyparsing!=0.17.0,!=0.17.1,!=0.17.2, >=0.14.0 in /usr/local/lib/python3.8/dist-packages (from jsonschema>=2.6->nbformat>=4.4->nbconvert) (3.1.0)
Requirement already satisfied: zipp>=3.1.0 in /usr/local/lib/python3.8/dist-packages (from importlib-resources>=1.4.0->jsonschema>=2.6->nbformat>=4.4->nbconvert) (3.13.0)
Reading package lists... Done
```

Building dependency tree
Reading state information... Done
The following package was automatically installed and is no longer required:
libnvidia-common-510

Use 'sudo apt autoremove' to remove it.

The following additional packages will be installed:

dvisvgm fonts-droid-fallback fonts-lato fonts-lmodern fonts-noto-mono
fonts-texgyre fonts-urw-base35 javascript-common libapache-pom-java
libcommons-logging-java libcommons-parent-java libfontbox-java libgs9
libgs9-common libidn11 libijs-0.35 libjbig2dec0 libjs-jquery libkpathsea6
libpdfbox-java libptexenc1 libruby2.7 libsynchronet2 libteckit0 libtexlua53
libtexluajit2 libzip-0.13 lmodern poppler-data preview-latex-style rake
ruby ruby-minitest ruby-net-telnet ruby-power-assert ruby-test-unit
ruby-xmllrpc ruby2.7 rubygems-integration tlutils teckit tex-common tex-gyre
texlive-base texlive-binaries texlive-latex-base texlive-latex-extra
texlive-latex-recommended texlive-pictures tipa xfonts-encodings
xfonts-utils

Suggested packages:

fonts-noto fonts-freefont-otf | fonts-freefont-ttf apache2 | lighttpd
| httpd libavalon-framework-java libcommons-logging-java-doc
libexcalibur-logkit-java liblog4j1.2-java poppler-utils ghostscript
fonts-japanese-mincho | fonts-ipafont-mincho fonts-japanese-gothic
| fonts-ipafont-gothic fonts-arphic-ukai fonts-arphic-uming fonts-nanum ri
ruby-dev bundler debhelper gv | postscript-viewer perl-tk xpdf | pdf-viewer
xzdec texlive-fonts-recommended-doc texlive-latex-base-doc python3-pygments
icc-profiles libfile-which-perl libspreadsheet-parseexcel-perl
texlive-latex-extra-doc texlive-latex-recommended-doc texlive-luatex
texlive-pstricks dot2tex prerex ruby-tcltk | libtcltk-ruby
texlive-pictures-doc vprerex

The following NEW packages will be installed:

dvisvgm fonts-droid-fallback fonts-lato fonts-lmodern fonts-noto-mono
fonts-texgyre fonts-urw-base35 javascript-common libapache-pom-java
libcommons-logging-java libcommons-parent-java libfontbox-java libgs9
libgs9-common libidn11 libijs-0.35 libjbig2dec0 libjs-jquery libkpathsea6
libpdfbox-java libptexenc1 libruby2.7 libsynchronet2 libteckit0 libtexlua53
libtexluajit2 libzip-0.13 lmodern poppler-data preview-latex-style rake
ruby ruby-minitest ruby-net-telnet ruby-power-assert ruby-test-unit
ruby-xmllrpc ruby2.7 rubygems-integration tlutils teckit tex-common tex-gyre
texlive-base texlive-binaries texlive-fonts-recommended texlive-latex-base
texlive-latex-extra texlive-latex-recommended texlive-pictures
texlive-plain-generic texlive-xetex tipa xfonts-encodings xfonts-utils

0 upgraded, 55 newly installed, 0 to remove and 21 not upgraded.

Need to get 169 MB of archives.

After this operation, 536 MB of additional disk space will be used.

Get:1 http://archive.ubuntu.com/ubuntu focal/main amd64 fonts-droid-fallback all 1:6.0.1r16-1.1 [1,805 kB]
Get:2 http://archive.ubuntu.com/ubuntu focal/main amd64 fonts-lato all 2.0-2 [2,698 kB]
Get:3 http://archive.ubuntu.com/ubuntu focal/main amd64 poppler-data all 0.4.9-2 [1,475 kB]
Get:4 http://archive.ubuntu.com/ubuntu focal/universe amd64 tex-common all 6.13 [32.7 kB]
Get:5 http://archive.ubuntu.com/ubuntu focal/main amd64 fonts-urw-base35 all 20170801.1-3 [6,333 kB]
Get:6 http://archive.ubuntu.com/ubuntu focal-updates/main amd64 libgs9-common all 9.50~dfsg-5ubuntu4.6 [681 kB]
Get:7 http://archive.ubuntu.com/ubuntu focal/main amd64 libidn11 amd64 1.33-2.2ubuntu2 [46.2 kB]
Get:8 http://archive.ubuntu.com/ubuntu focal/main amd64 libijs-0.35 amd64 0.35-15 [15.7 kB]
Get:9 http://archive.ubuntu.com/ubuntu focal/main amd64 libjbig2dec0 amd64 0.18-1ubuntu1 [60.0 kB]
Get:10 http://archive.ubuntu.com/ubuntu focal-updates/main amd64 libgs9 amd64 9.50~dfsg-5ubuntu4.6 [2,173 kB]
Get:11 http://archive.ubuntu.com/ubuntu focal/main amd64 libkpathsea6 amd64 2019.20190605.51237-3build2 [57.0 kB]
Get:12 http://archive.ubuntu.com/ubuntu focal/universe amd64 dvisvgm amd64 2.8.1-1build1 [1,048 kB]
Get:13 http://archive.ubuntu.com/ubuntu focal/universe amd64 fonts-lmodern all 2.004.5-6 [4,532 kB]
Get:14 http://archive.ubuntu.com/ubuntu focal-updates/main amd64 fonts-noto-mono all 20200323-1build1~ubuntu20.04.1 [80.6 kB]
Get:15 http://archive.ubuntu.com/ubuntu focal/universe amd64 fonts-texgyre all 20180621-3 [10.2 MB]
Get:16 http://archive.ubuntu.com/ubuntu focal/main amd64 javascript-common all 11 [6,066 B]
Get:17 http://archive.ubuntu.com/ubuntu focal/universe amd64 libapache-pom-java all 18-1 [4,720 B]
Get:18 http://archive.ubuntu.com/ubuntu focal/universe amd64 libcommons-parent-java all 43-1 [10.8 kB]
Get:19 http://archive.ubuntu.com/ubuntu focal/universe amd64 libcommons-logging-java all 1.2-2 [60.3 kB]
Get:20 http://archive.ubuntu.com/ubuntu focal/main amd64 libjs-jquery all 3.3.1~dfsg-3 [329 kB]
Get:21 http://archive.ubuntu.com/ubuntu focal/main amd64 libptexenc1 amd64 2019.20190605.51237-3build2 [35.5 kB]
Get:22 http://archive.ubuntu.com/ubuntu focal/main amd64 rubygems-integration all 1.16 [5,092 B]
Get:23 http://archive.ubuntu.com/ubuntu focal-updates/main amd64 ruby2.7 amd64 2.7.0-5ubuntu1.7 [95.6 kB]
Get:24 http://archive.ubuntu.com/ubuntu focal/main amd64 ruby amd64 1:2.7+1 [5,412 B]
Get:25 http://archive.ubuntu.com/ubuntu focal/main amd64 rake all 13.0.1-4 [61.6 kB]
Get:26 http://archive.ubuntu.com/ubuntu focal/main amd64 ruby-minitest all 5.13.0-1 [40.9 kB]
Get:27 http://archive.ubuntu.com/ubuntu focal/main amd64 ruby-net-telnet all 0.1.1-2 [12.6 kB]
Get:28 http://archive.ubuntu.com/ubuntu focal/main amd64 ruby-power-assert all 1.1.7-1 [11.4 kB]
Get:29 http://archive.ubuntu.com/ubuntu focal/main amd64 ruby-test-unit all 3.3.5-1 [73.2 kB]
Get:30 http://archive.ubuntu.com/ubuntu focal/main amd64 ruby-xmllrpc all 0.3.0-2 [23.8 kB]
Get:31 http://archive.ubuntu.com/ubuntu focal-updates/main amd64 libruby2.7 amd64 2.7.0-5ubuntu1.7 [3,533 kB]
Get:32 http://archive.ubuntu.com/ubuntu focal/main amd64 libsynchronet2 amd64 2019.20190605.51237-3build2 [55.0 kB]
Get:33 http://archive.ubuntu.com/ubuntu focal/universe amd64 libteckit0 amd64 2.5.8+ds2-5ubuntu2 [320 kB]
Get:34 http://archive.ubuntu.com/ubuntu focal/main amd64 libtexlua53 amd64 2019.20190605.51237-3build2 [105 kB]
Get:35 http://archive.ubuntu.com/ubuntu focal/main amd64 libtexluajit2 amd64 2019.20190605.51237-3build2 [235 kB]

```

B]
Get:36 http://archive.ubuntu.com/ubuntu focal/universe amd64 libzzip-0-13 amd64 0.13.62-3.2ubuntu1 [26.2 kB]
Get:37 http://archive.ubuntu.com/ubuntu focal/main amd64 xfonts-encodings all 1:1.0.5-0ubuntu1 [573 kB]
Get:38 http://archive.ubuntu.com/ubuntu focal/main amd64 xfonts-utils amd64 1:7.7+6 [91.5 kB]
Get:39 http://archive.ubuntu.com/ubuntu focal/universe amd64 lmodern all 2.004.5-6 [9,474 kB]
Get:40 http://archive.ubuntu.com/ubuntu focal/universe amd64 preview-latex-style all 11.91-2ubuntu2 [184 kB]
Get:41 http://archive.ubuntu.com/ubuntu focal/main amd64 tlutils amd64 1.41-3 [56.1 kB]
Get:42 http://archive.ubuntu.com/ubuntu focal/universe amd64 teckit amd64 2.5.8+ds2-5ubuntu2 [687 kB]
Get:43 http://archive.ubuntu.com/ubuntu focal/universe amd64 tex-gyre all 20180621-3 [6,209 kB]
Get:44 http://archive.ubuntu.com/ubuntu focal/universe amd64 texlive-binaries amd64 2019.20190605.51237-3build2 [8,041 kB]
Get:45 http://archive.ubuntu.com/ubuntu focal/universe amd64 texlive-base all 2019.20200218-1 [20.8 MB]
Get:46 http://archive.ubuntu.com/ubuntu focal/universe amd64 texlive-fonts-recommended all 2019.20200218-1 [4,972 kB]
Get:47 http://archive.ubuntu.com/ubuntu focal/universe amd64 texlive-latex-base all 2019.20200218-1 [990 kB]
Get:48 http://archive.ubuntu.com/ubuntu focal/universe amd64 libfontbox-java all 1:1.8.16-2 [207 kB]
Get:49 http://archive.ubuntu.com/ubuntu focal/universe amd64 libpdfbox-java all 1:1.8.16-2 [5,199 kB]
Get:50 http://archive.ubuntu.com/ubuntu focal/universe amd64 texlive-latex-recommended all 2019.20200218-1 [15.7 MB]
Get:51 http://archive.ubuntu.com/ubuntu focal/universe amd64 texlive-pictures all 2019.20200218-1 [4,492 kB]
Get:52 http://archive.ubuntu.com/ubuntu focal/universe amd64 texlive-latex-extra all 2019.20200218-1 [12.5 MB]
Get:53 http://archive.ubuntu.com/ubuntu focal/universe amd64 texlive-plain-generic all 2019.20200218-1 [24.6 MB]
B]
Get:54 http://archive.ubuntu.com/ubuntu focal/universe amd64 tipa all 2:1.3-20 [2,978 kB]
Get:55 http://archive.ubuntu.com/ubuntu focal/universe amd64 texlive-xetex all 2019.20200218-1 [14.6 MB]
Fetched 169 MB in 25s (6,800 kB/s)
debconf: unable to initialize frontend: Dialog
debconf: (No usable dialog-like program is installed, so the dialog based frontend cannot be used. at /usr/share/perl5/Debconf/FrontEnd/Dialog.pm line 76, <> line 55.)
debconf: falling back to frontend: Readline
debconf: unable to initialize frontend: Readline
debconf: (This frontend requires a controlling tty.)
debconf: falling back to frontend: Teletype
dpkg-preconfigure: unable to re-open stdin:
Selecting previously unselected package fonts-droid-fallback.
(Reading database ... 128126 files and directories currently installed.)
Preparing to unpack .../00-fonts-droid-fallback_1%3a6.0.1r16-1.1_all.deb ...
Unpacking fonts-droid-fallback (1:6.0.1r16-1.1) ...
Selecting previously unselected package fonts-lato.
Preparing to unpack .../01-fonts-lato_2.0-2_all.deb ...
Unpacking fonts-lato (2.0-2) ...
Selecting previously unselected package poppler-data.
Preparing to unpack .../02-poppler-data_0.4.9-2_all.deb ...
Unpacking poppler-data (0.4.9-2) ...
Selecting previously unselected package tex-common.
Preparing to unpack .../03-tex-common_6.13_all.deb ...
Unpacking tex-common (6.13) ...
Selecting previously unselected package fonts-urw-base35.
Preparing to unpack .../04-fonts-urw-base35_20170801.1-3_all.deb ...
Unpacking fonts-urw-base35 (20170801.1-3) ...
Selecting previously unselected package libgs9-common.
Preparing to unpack .../05-libgs9-common_9.50~dfsg-5ubuntu4.6_all.deb ...
Unpacking libgs9-common (9.50~dfsg-5ubuntu4.6) ...
Selecting previously unselected package libidn11:amd64.
Preparing to unpack .../06-libidn11_1.33-2.2ubuntu2_amd64.deb ...
Unpacking libidn11:amd64 (1.33-2.2ubuntu2) ...
Selecting previously unselected package libijs-0.35:amd64.
Preparing to unpack .../07-libijs-0.35_0.35-15_amd64.deb ...
Unpacking libijs-0.35:amd64 (0.35-15) ...
Selecting previously unselected package libjbig2dec0:amd64.
Preparing to unpack .../08-libjbig2dec0_0.18-1ubuntu1_amd64.deb ...
Unpacking libjbig2dec0:amd64 (0.18-1ubuntu1) ...
Selecting previously unselected package libgs9:amd64.
Preparing to unpack .../09-libgs9_9.50~dfsg-5ubuntu4.6_amd64.deb ...
Unpacking libgs9:amd64 (9.50~dfsg-5ubuntu4.6) ...
Selecting previously unselected package libkpathsea6:amd64.
Preparing to unpack .../10-libkpathsea6_2019.20190605.51237-3build2_amd64.deb ...
Unpacking libkpathsea6:amd64 (2019.20190605.51237-3build2) ...
Selecting previously unselected package dvisvgm.
Preparing to unpack .../11-dvisvgm_2.8.1-1build1_amd64.deb ...
Unpacking dvisvgm (2.8.1-1build1) ...
Selecting previously unselected package fonts-lmodern.
Preparing to unpack .../12-fonts-lmodern_2.004.5-6_all.deb ...
Unpacking fonts-lmodern (2.004.5-6) ...
Selecting previously unselected package fonts-noto-mono.
Preparing to unpack .../13-fonts-noto-mono_20200323-1build1~ubuntu20.04.1_all.deb ...
Unpacking fonts-noto-mono (20200323-1build1~ubuntu20.04.1) ...
Selecting previously unselected package fonts-texgyre.
Preparing to unpack .../14-fonts-texgyre_20180621-3_all.deb ...
Unpacking fonts-texgyre (20180621-3) ...
Selecting previously unselected package javascript-common.
Preparing to unpack .../15-javascript-common_11_all.deb ...
Unpacking javascript-common (11) ...

```

Selecting previously unselected package libapache-pom-java.
Preparing to unpack .../16-libapache-pom-java_18-1_all.deb ...
Unpacking libapache-pom-java (18-1) ...
Selecting previously unselected package libcommons-parent-java.
Preparing to unpack .../17-libcommons-parent-java_43-1_all.deb ...
Unpacking libcommons-parent-java (43-1) ...
Selecting previously unselected package libcommons-logging-java.
Preparing to unpack .../18-libcommons-logging-java_1.2-2_all.deb ...
Unpacking libcommons-logging-java (1.2-2) ...
Selecting previously unselected package libjs-jquery.
Preparing to unpack .../19-libjs-jquery_3.3.1-dfsg-3_all.deb ...
Unpacking libjs-jquery (3.3.1-dfsg-3) ...
Selecting previously unselected package libptexenc1:amd64.
Preparing to unpack .../20-libptexenc1_2019.20190605.51237-3build2_amd64.deb ...
Unpacking libptexenc1:amd64 (2019.20190605.51237-3build2) ...
Selecting previously unselected package rubygems-integration.
Preparing to unpack .../21-rubygems-integration_1.16_all.deb ...
Unpacking rubygems-integration (1.16) ...
Selecting previously unselected package ruby2.7.
Preparing to unpack .../22-ruby2.7_2.7.0-5ubuntu1.7_amd64.deb ...
Unpacking ruby2.7 (2.7.0-5ubuntu1.7) ...
Selecting previously unselected package ruby.
Preparing to unpack .../23-ruby_1%3a2.7+1_amd64.deb ...
Unpacking ruby (1:2.7+1) ...
Selecting previously unselected package rake.
Preparing to unpack .../24-rake_13.0.1-4_all.deb ...
Unpacking rake (13.0.1-4) ...
Selecting previously unselected package ruby-minitest.
Preparing to unpack .../25-ruby-minitest_5.13.0-1_all.deb ...
Unpacking ruby-minitest (5.13.0-1) ...
Selecting previously unselected package ruby-net-telnet.
Preparing to unpack .../26-ruby-net-telnet_0.1.1-2_all.deb ...
Unpacking ruby-net-telnet (0.1.1-2) ...
Selecting previously unselected package ruby-power-assert.
Preparing to unpack .../27-ruby-power-assert_1.1.7-1_all.deb ...
Unpacking ruby-power-assert (1.1.7-1) ...
Selecting previously unselected package ruby-test-unit.
Preparing to unpack .../28-ruby-test-unit_3.3.5-1_all.deb ...
Unpacking ruby-test-unit (3.3.5-1) ...
Selecting previously unselected package ruby-xmlrpc.
Preparing to unpack .../29-ruby-xmlrpc_0.3.0-2_all.deb ...
Unpacking ruby-xmlrpc (0.3.0-2) ...
Selecting previously unselected package libruby2.7:amd64.
Preparing to unpack .../30-libruby2.7_2.7.0-5ubuntu1.7_amd64.deb ...
Unpacking libruby2.7:amd64 (2.7.0-5ubuntu1.7) ...
Selecting previously unselected package libsyntax2:amd64.
Preparing to unpack .../31-libsyntax2_2019.20190605.51237-3build2_amd64.deb ...
Unpacking libsyntax2:amd64 (2019.20190605.51237-3build2) ...
Selecting previously unselected package libteckit0:amd64.
Preparing to unpack .../32-libteckit0_2.5.8+ds2-5ubuntu2_amd64.deb ...
Unpacking libteckit0:amd64 (2.5.8+ds2-5ubuntu2) ...
Selecting previously unselected package libtexlua53:amd64.
Preparing to unpack .../33-libtexlua53_2019.20190605.51237-3build2_amd64.deb ...
Unpacking libtexlua53:amd64 (2019.20190605.51237-3build2) ...
Selecting previously unselected package libtexluajit2:amd64.
Preparing to unpack .../34-libtexluajit2_2019.20190605.51237-3build2_amd64.deb ...
Unpacking libtexluajit2:amd64 (2019.20190605.51237-3build2) ...
Selecting previously unselected package libzip-0-13:amd64.
Preparing to unpack .../35-libzip-0-13_0.13.62-3.2ubuntu1_amd64.deb ...
Unpacking libzip-0-13:amd64 (0.13.62-3.2ubuntu1) ...
Selecting previously unselected package xfonts-encodings.
Preparing to unpack .../36-xfonts-encodings_1%3a1.0.5-0ubuntu1_all.deb ...
Unpacking xfonts-encodings (1:1.0.5-0ubuntu1) ...
Selecting previously unselected package xfonts-utils.
Preparing to unpack .../37-xfonts-utils_1%3a7.7+6_amd64.deb ...
Unpacking xfonts-utils (1:7.7+6) ...
Selecting previously unselected package lmodern.
Preparing to unpack .../38-lmodern_2.004.5-6_all.deb ...
Unpacking lmodern (2.004.5-6) ...
Selecting previously unselected package preview-latex-style.
Preparing to unpack .../39-preview-latex-style_11.91-2ubuntu2_all.deb ...
Unpacking preview-latex-style (11.91-2ubuntu2) ...
Selecting previously unselected package tlutils.
Preparing to unpack .../40-tlutils_1.41-3_amd64.deb ...
Unpacking tlutils (1.41-3) ...
Selecting previously unselected package teckit.
Preparing to unpack .../41-teckit_2.5.8+ds2-5ubuntu2_amd64.deb ...
Unpacking teckit (2.5.8+ds2-5ubuntu2) ...
Selecting previously unselected package tex-gyre.
Preparing to unpack .../42-tex-gyre_20180621-3_all.deb ...
Unpacking tex-gyre (20180621-3) ...
Selecting previously unselected package texlive-binaries.
Preparing to unpack .../43-texlive-binaries_2019.20190605.51237-3build2_amd64.deb ...

```

Unpacking texlive-binaries (2019.20190605.51237-3build2) ...
Selecting previously unselected package texlive-base.
Preparing to unpack .../44-texlive-base_2019.20200218-1_all.deb ...
Unpacking texlive-base (2019.20200218-1) ...
Selecting previously unselected package texlive-fonts-recommended.
Preparing to unpack .../45-texlive-fonts-recommended_2019.20200218-1_all.deb ...
Unpacking texlive-fonts-recommended (2019.20200218-1) ...
Selecting previously unselected package texlive-latex-base.
Preparing to unpack .../46-texlive-latex-base_2019.20200218-1_all.deb ...
Unpacking texlive-latex-base (2019.20200218-1) ...
Selecting previously unselected package libfontbox-java.
Preparing to unpack .../47-libfontbox-java_1%3a1.8.16-2_all.deb ...
Unpacking libfontbox-java (1:1.8.16-2) ...
Selecting previously unselected package libpdfbox-java.
Preparing to unpack .../48-libpdfbox-java_1%3a1.8.16-2_all.deb ...
Unpacking libpdfbox-java (1:1.8.16-2) ...
Selecting previously unselected package texlive-latex-recommended.
Preparing to unpack .../49-texlive-latex-recommended_2019.20200218-1_all.deb ...
Unpacking texlive-latex-recommended (2019.20200218-1) ...
Selecting previously unselected package texlive-pictures.
Preparing to unpack .../50-texlive-pictures_2019.20200218-1_all.deb ...
Unpacking texlive-pictures (2019.20200218-1) ...
Selecting previously unselected package texlive-latex-extra.
Preparing to unpack .../51-texlive-latex-extra_2019.20200218-1_all.deb ...
Unpacking texlive-latex-extra (2019.20200218-1) ...
Selecting previously unselected package texlive-plain-generic.
Preparing to unpack .../52-texlive-plain-generic_2019.20200218-1_all.deb ...
Unpacking texlive-plain-generic (2019.20200218-1) ...
Selecting previously unselected package tipa.
Preparing to unpack .../53-tipa_2%3a1.3-20_all.deb ...
Unpacking tipa (2:1.3-20) ...
Selecting previously unselected package texlive-xetex.
Preparing to unpack .../54-texlive-xetex_2019.20200218-1_all.deb ...
Unpacking texlive-xetex (2019.20200218-1) ...
Setting up javascript-common (11) ...
Setting up fonts-lato (2.0-2) ...
Setting up fonts-noto-mono (20200323-1build1-ubuntu20.04.1) ...
Setting up ruby-power-assert (1.1.7-1) ...
Setting up libtexlua53:amd64 (2019.20190605.51237-3build2) ...
Setting up libijs-0.35:amd64 (0.35-15) ...
Setting up libtexluajit2:amd64 (2019.20190605.51237-3build2) ...
Setting up libfontbox-java (1:1.8.16-2) ...
Setting up rubygems-integration (1.16) ...
Setting up libzzip-0-13:amd64 (0.13.62-3.2ubuntu1) ...
Setting up fonts-urw-base35 (20170801.1-3) ...
Setting up poppler-data (0.4.9-2) ...
Setting up ruby-minitest (5.13.0-1) ...
Setting up tex-common (6.13) ...
debconf: unable to initialize frontend: Dialog
debconf: (No usable dialog-like program is installed, so the dialog based frontend cannot be used. at /usr/share/perl5/Debconf/FrontEnd/Dialog.pm line 76.)
debconf: falling back to frontend: Readline
update-language: texlive-base not installed and configured, doing nothing!
Setting up ruby-test-unit (3.3.5-1) ...
Setting up libjbig2dec0:amd64 (0.18-1ubuntu1) ...
Setting up libidn11:amd64 (1.33-2.2ubuntu2) ...
Setting up libteckit0:amd64 (2.5.8+ds2-5ubuntu2) ...
Setting up libapache-pom-java (18-1) ...
Setting up ruby-net-telnet (0.1.1-2) ...
Setting up xfonts-encodings (1:1.0.5-0ubuntu1) ...
Setting up tlutils (1.41-3) ...
Setting up fonts-texgyre (20180621-3) ...
Setting up libkpathsea6:amd64 (2019.20190605.51237-3build2) ...
Setting up fonts-lmodern (2.004.5-6) ...
Setting up fonts-droid-fallback (1:6.0.1r16-1.1) ...
Setting up libjs-jquery (3.3.1~dfsg-3) ...
Setting up ruby-xmlrpc (0.3.0-2) ...
Setting up libsynchronet2:amd64 (2019.20190605.51237-3build2) ...
Setting up libgs9-common (9.50~dfsg-5ubuntu4.6) ...
Setting up teckit (2.5.8+ds2-5ubuntu2) ...
Setting up libpdfbox-java (1:1.8.16-2) ...
Setting up libgs9:amd64 (9.50~dfsg-5ubuntu4.6) ...
Setting up preview-latex-style (11.91-2ubuntu2) ...
Setting up libcommons-parent-java (43-1) ...
Setting up dvisvgm (2.8.1-1build1) ...
Setting up libcommons-logging-java (1.2-2) ...
Setting up xfonts-utils (1:7.7+6) ...
Setting up libptexenc1:amd64 (2019.20190605.51237-3build2) ...
Setting up texlive-binaries (2019.20190605.51237-3build2) ...
update-alternatives: using /usr/bin/xdvi-xaw to provide /usr/bin/xdvi.bin (xdvi.bin) in auto mode
update-alternatives: using /usr/bin/bibtex.original to provide /usr/bin/bibtex (bibtex) in auto mode
Setting up lmodern (2.004.5-6) ...
Setting up texlive-base (2019.20200218-1) ...

```

```

mktexlsr: Updating /var/lib/texmf/ls-R-TEXLIVEDIST...
mktexlsr: Updating /var/lib/texmf/ls-R-TEXMFMAIN...
mktexlsr: Updating /var/lib/texmf/ls-R...
mktexlsr: Done.
tl-paper: setting paper size for dvips to a4: /var/lib/texmf/dvips/config/config-paper.ps
tl-paper: setting paper size for dvipdfmx to a4: /var/lib/texmf/dvipdfmx/dvipdfmx-paper.cfg
tl-paper: setting paper size for xdvi to a4: /var/lib/texmf/xdvi/XDvi-paper
tl-paper: setting paper size for pdftex to a4: /var/lib/texmf/tex/generic/config/pdftexconfig.tex
debconf: unable to initialize frontend: Dialog
debconf: (No usable dialog-like program is installed, so the dialog based frontend cannot be used. at /usr/share/perl5/Debconf/FrontEnd/Dialog.pm line 76.)
debconf: falling back to frontend: Readline
Setting up tex-gyre (20180621-3) ...
Setting up texlive-plain-generic (2019.20200218-1) ...
Setting up texlive-latex-base (2019.20200218-1) ...
Setting up texlive-latex-recommended (2019.20200218-1) ...
Setting up texlive-pictures (2019.20200218-1) ...
Setting up texlive-fonts-recommended (2019.20200218-1) ...
Setting up tipa (2:1.3-20) ...
Regenerating '/var/lib/texmf/fmtutil.cnf-DEBIAN'... done.
Regenerating '/var/lib/texmf/fmtutil.cnf-TEXLIVEDIST'... done.
update-fmtutil has updated the following file(s):
    /var/lib/texmf/fmtutil.cnf-DEBIAN
    /var/lib/texmf/fmtutil.cnf-TEXLIVEDIST
If you want to activate the changes in the above file(s),
you should run fmtutil-sys or fmtutil.
Setting up texlive-latex-extra (2019.20200218-1) ...
Setting up texlive-xetex (2019.20200218-1) ...
Setting up rake (13.0.1-4) ...
Setting up libruby2.7:amd64 (2.7.0-5ubuntu1.7) ...
Setting up ruby2.7 (2.7.0-5ubuntu1.7) ...
Setting up ruby (1:2.7+1) ...
Processing triggers for fontconfig (2.13.1-2ubuntu3) ...
Processing triggers for mime-support (3.64ubuntu1) ...
Processing triggers for libc-bin (2.31-0ubuntu9.9) ...
Processing triggers for man-db (2.9.1-1) ...
Processing triggers for tex-common (6.13) ...
debconf: unable to initialize frontend: Dialog
debconf: (No usable dialog-like program is installed, so the dialog based frontend cannot be used. at /usr/share/perl5/Debconf/FrontEnd/Dialog.pm line 76.)
debconf: falling back to frontend: Readline
Running updmap-sys. This may take some time... done.
Running mktexlsr /var/lib/texmf ... done.
Building format(s) --all.
    This may take some time... done.

```

```

In [23]: !jupyter nbconvert --to html "/content/drive/MyDrive/CS6700/CS6700_Tutorial_4_QLearning_SARSA/CS6700_Tutorial_4_QLearning_SARSA.ipynb"
[NbConvertApp] Converting notebook /content/drive/MyDrive/CS6700/CS6700_Tutorial_4_QLearning_SARSA/CS6700_Tutorial_4_QLearning_SARSA.ipynb to html
[NbConvertApp] Writing 790082 bytes to /content/drive/MyDrive/CS6700/CS6700_Tutorial_4_QLearning_SARSA/CS6700_Tutorial_4_QLearning_SARSA.html

```

In []:

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js