ENPM808F - Robot Learning
Homework 4 - Dots and Boxes

Rama Prashanth (116428941)

This project leverages the reinforcement learning algorithm of Q learning to implement the dots and boxes game. The Q learning algorithm is basically an agent which learns to act by interacting with its environment by experiencing rewards by taking actions. In the game of Dots and Boxes, the players take turns drawing lines and when a player completes a box with four lines they get a point and get to draw again. The game continues until all lines are filled and player with most points wins the game.

Rewards:

- Completion of box +1
- Win +5
- Drawing line 0

Metrics:

- Win rate (number of games Q learning agent can win)
- Loss rate
- Draw rate

Q-Learning essentially involves finding the Q-values represented by Q(S,a) of the state-action pairs and given as follows.

$$Q(s_t, a_t) \leftarrow \underbrace{Q(s_t, a_t)}_{\text{old value}} + \underbrace{\alpha_t}_{\text{learning rate}} \cdot \left(\overbrace{\underbrace{r_{t+1}}_{\text{reward}} + \underbrace{\gamma}_{\text{discount factor}} \cdot \underbrace{\max_a Q(s_{t+1}, a)}_{\text{estimate of optimal future value}}}^{\text{learned value}} - \underbrace{Q(s_t, a_t)}_{\text{old value}} \right)$$

Q Table - Epsilon Greedy Policy Parameters

epsilon = 0.05
Gamma = 0.6
Alpha = 1e-4

The self-play learning phase was iterated over 100, 1000 and 10,000 games and results are as follows.

```
Starting at game 1
Game 100 Test Results
Current win % over agent random: 43.00%
Current loss % over agent random: 39.00%
Current draw % over agent random: 18.00%
100,random,0.43,0.18,0.39

Game 1000 Test Results
Current win % over agent random: 44.90%
Current loss % over agent random: 40.60%
Current draw % over agent random: 14.50%
1000,random,0.449,0.145,0.406

Game 10000 Test Results
Current win % over agent random: 43.50%
Current loss % over agent random: 41.39%
Current draw % over agent random: 15.11%
10000,random,0.435,0.1511,0.4139
```

*2*2 QTable Learning*

```
Starting at game 1
Game 100 Test Results
Current win % over agent random: 47.00%
Current loss % over agent random: 53.00%
Current draw % over agent random: 0.00%
100,random,0.47,0.0,0.53

Game 1000 Test Results
Current win % over agent random: 49.80%
Current loss % over agent random: 50.20%
Current draw % over agent random: 0.00%
1000,random,0.498,0.0,0.502

Game 10000 Test Results
Current win % over agent random: 53.22%
Current loss % over agent random: 46.78%
Current draw % over agent random: 0.00%
10000,random,0.5322,0.0,0.4678
```

*3*3 QTable Learning*

DQN uses Q-learning along with deep neural nets as a function approximation to the Q-values so that it is not mandatory for every state action pair to be explored. Also, in DQN the Q-Function is not updated online but rather records of gameplay are stores as replay table which acts as experience relay which is later randomly sampled to update Q-Function.

```
Starting Game at 1
Game 100 Test Results
Current win % over agent random: 37.00%
Current loss % over agent random: 49.00%
Current draw % over agent random: 14.00%
100,random,0.37,0.14,0.49

Game 1000 Test Results
Current win % over agent random: 41.90%
Current loss % over agent random: 42.20%
Current draw % over agent random: 15.90%
1000,random,0.419,0.159,0.422

Game 10000 Test Results
Current win % over agent random: 43.42%
Current loss % over agent random: 40.50%
Current draw % over agent random: 16.08%
10000,random,0.4342,0.1608,0.405

Finished!
```

*2*2 DQN Learning*

```
Starting Game at 1
Game 100 Test Results
Current win % over agent random: 44.00%
Current loss % over agent random: 56.00%
Current draw % over agent random: 0.00%
100,random,0.44,0.0,0.56

Game 1000 Test Results
Current win % over agent random: 55.90%
Current loss % over agent random: 44.10%
Current draw % over agent random: 0.00%
1000,random,0.559,0.0,0.441

Game 10000 Test Results
Current win % over agent random: 53.60%
Current loss % over agent random: 46.40%
Current draw % over agent random: 0.00%
10000,random,0.536,0.0,0.464

Finished!
```

*3*3 DQN Learning*