

ENPM 808F: Robot Learning

Project 4

Amrish Baskaran

116301046

Report

Q-learning applied to the game Dots and Boxes

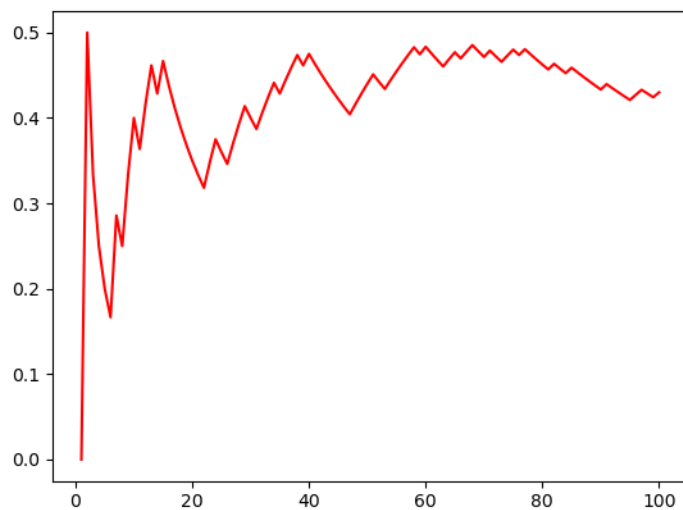
Rules provided - Players take turns drawing lines to complete boxes. A player is allowed to take another turn after completing a box. The game concludes when no more lines can be drawn and all boxes are claimed. Begin with a 2x2 grid. Assign a reward of +1 for completion of a box, +5 for a win, and 0 otherwise.

Training Self Play for 2x2 grid-

Both agents that are playing the games are using the same Q table to learn. The following are the results obtained

1. Training for 100 games-

Graph – Games vs Win percentage

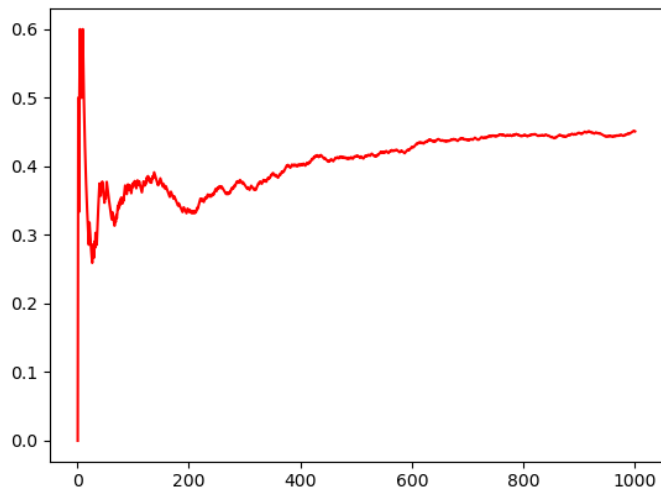


Wins-

Agent 1 = 41

Agent 2 = 59

2. Training for 1000 games-
Graph – Games vs Win percentage

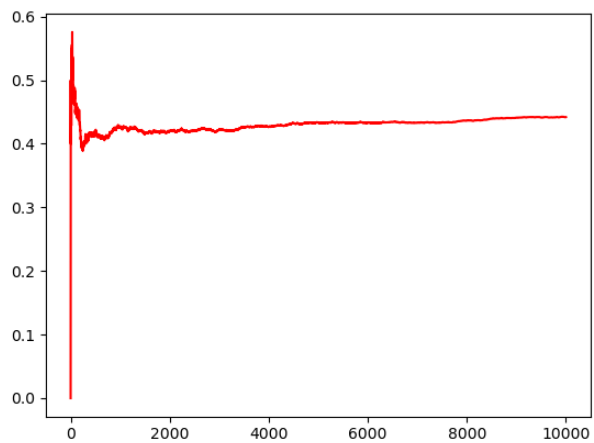


Wins-

Agent 1 = 47

Agent 2 = 53

3. Training for 10000 games-
Graph – Games vs Win percentage



Wins-

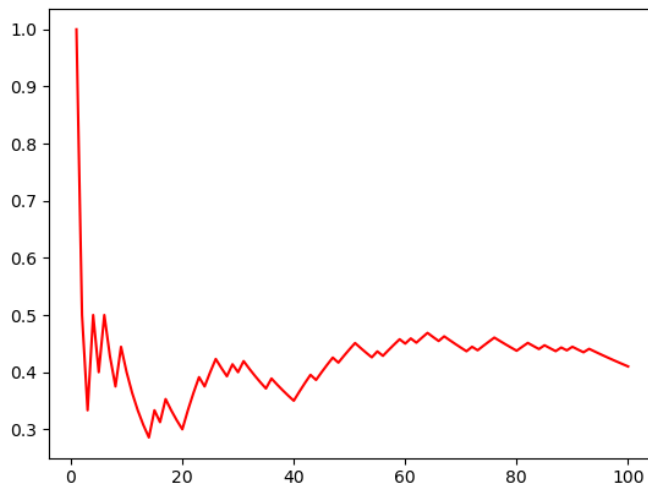
Agent 1 = 45

Agent 2 = 55

Testing these Q tables with a random agent for 2x2 grid-

One agent is playing with the Q table that is trained in the previous steps choosing the best action possible- that is Greedy Policy. And the other agent is playing random moves. Total number of games for testing 100 games.

1. Q table trained with 100 games

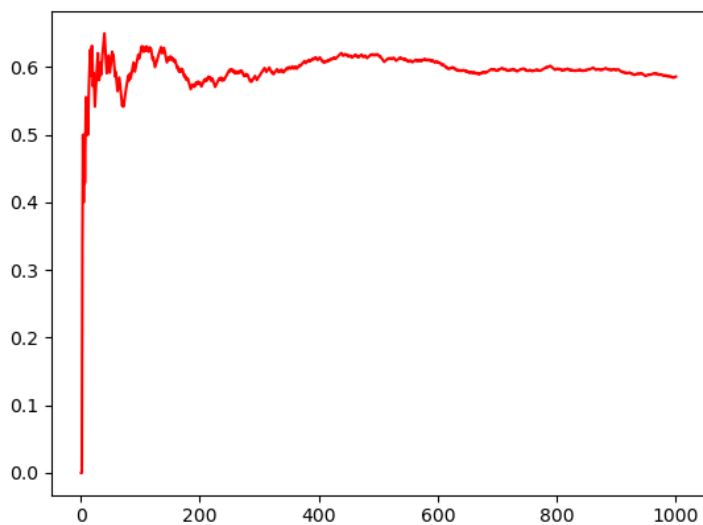


Wins-

Agent 1 = 40

Agent 2 = 60

2. Q table trained with 1000 games

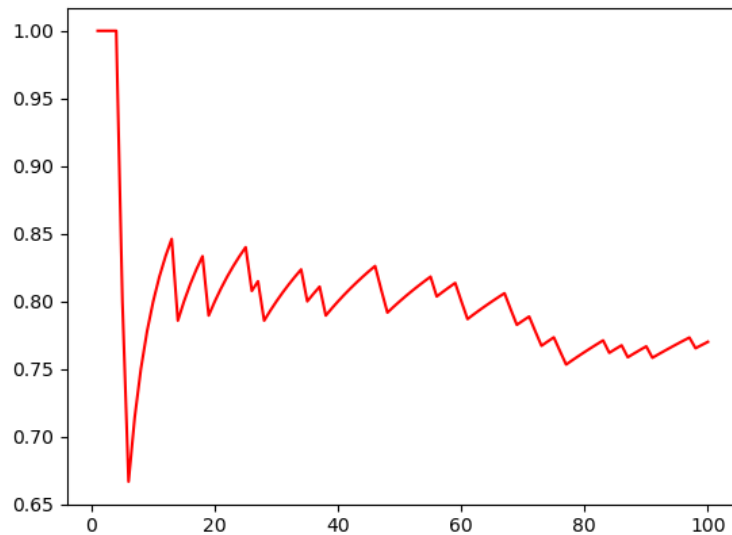


Wins-

Agent 1 = 58

Agent 2 = 42

3. Q table trained with 10000 games



Wins-

Agent 1 = 77

Agent 2 = 23

Potential for seeding Q table to 3x3 grid:

As the game is setup using states as binary encoding it is difficult to use the Q table from 2x2 grid to 3x3 grid.

If its required then the conversion from both must be determined and the appropriate values must be used be copied.

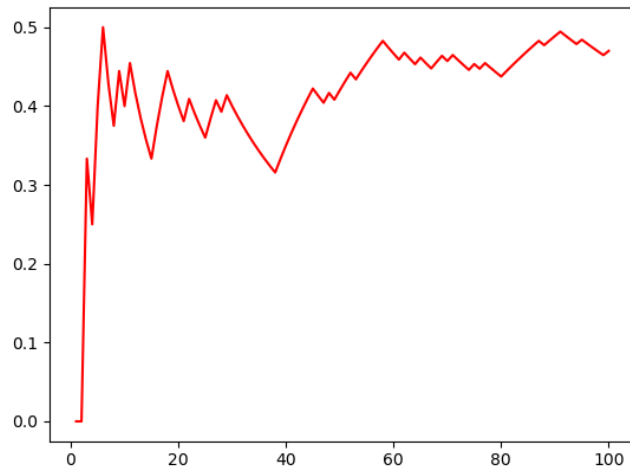
Another option is to use a different way to encode the states so that it's a part of 3x3.

Training Self Play for 3x3 grid-

Both agents that are playing the games are using the same Q table to learn. The following are the results obtained

1. Training for 100 games-

Graph – Games vs Win percentage



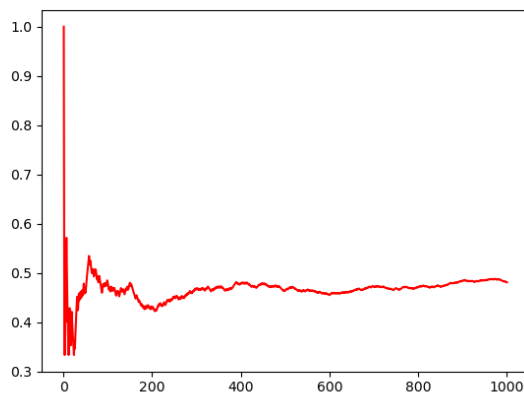
Wins-

Agent 1 = 48

Agent 2 = 52

2. Training for 1000 games-

Graph – Games vs Win percentage

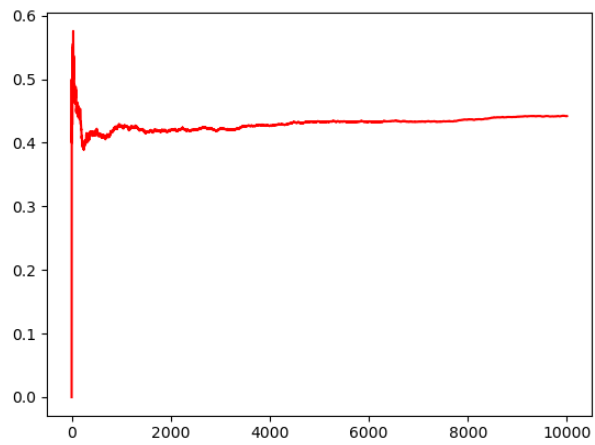


Wins-

Agent 1 = 47

Agent 2 = 53

3. Training for 10000 games-
Graph – Games vs Win percentage



Wins-

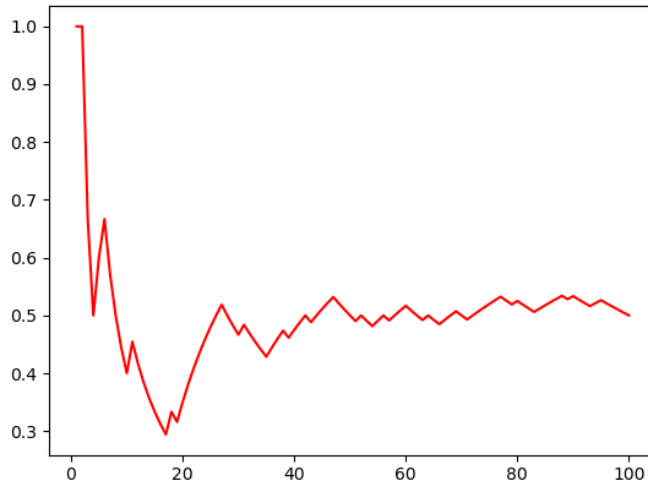
Agent 1 = 43

Agent 2 = 57

Testing these Q tables with a random agent for 3x3 grid-

One agent is playing with the Q table that is trained in the previous steps choosing the best action possible- that is Greedy Policy. And the other agent is playing random moves. Total number of games for testing 100 games.

1. Q table trained with 100 games

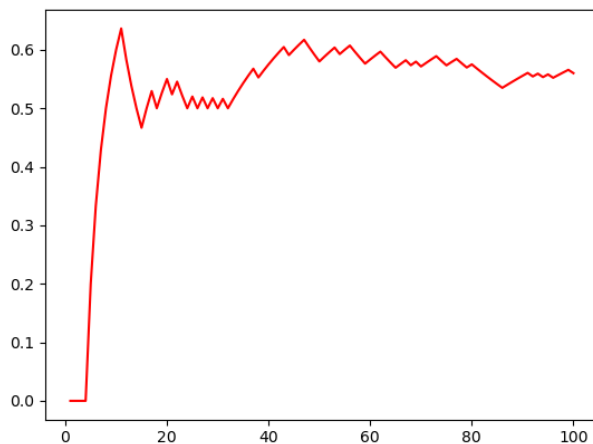


Wins-

Agent 1 = 50

Agent 2 = 50

2. Q table trained with 1000 games

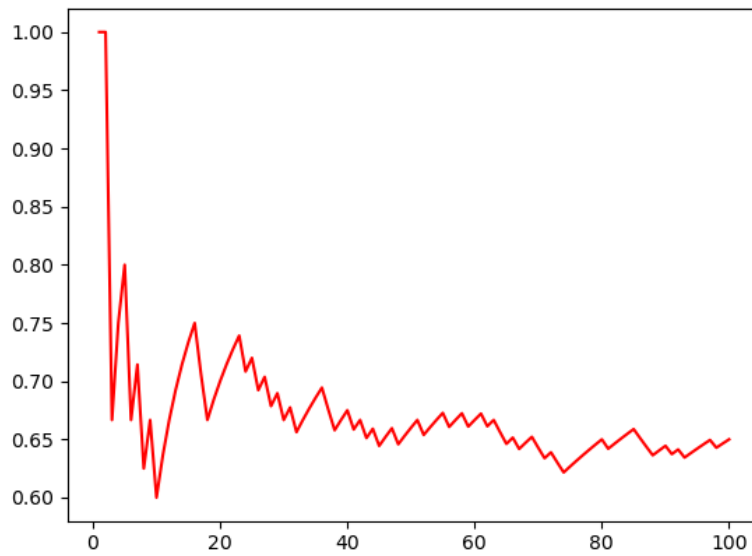


Wins-

Agent 1 = 58

Agent 2 = 42

3. Q table trained with 10000 games



Wins-

Agent 1 = 67

Agent 2 = 33