

Q-Learning Tic Tac Toe

Gnyana Teja Samudrala

April 2, 2019

Implementation

The system learns the game by playing against the other Q learning player. The player who makes the first move is selected at random. There will be two training sessions, one for the bot playing as 'X' and the other for the bot playing as 'O' against the human. So when the Q table is made for the bot playing as 'X', the other player in the training session will have a higher exploring rate, which makes it more probable to make moves at random. This is done so that the other bot learns all the states possible. If both players in the training session have same exploring rate then the bot doesnot learn few moves.

The plots of the number of wins, loses, ties and invalid moves over the number of episodes is given in the below figures. From the graphs we can observe that as the number of episodes increases, loses and ties saturate. On the other hand wins and ties keep increasing proportionally. The plots are in the order of wins, loses, ties and invalid moves.

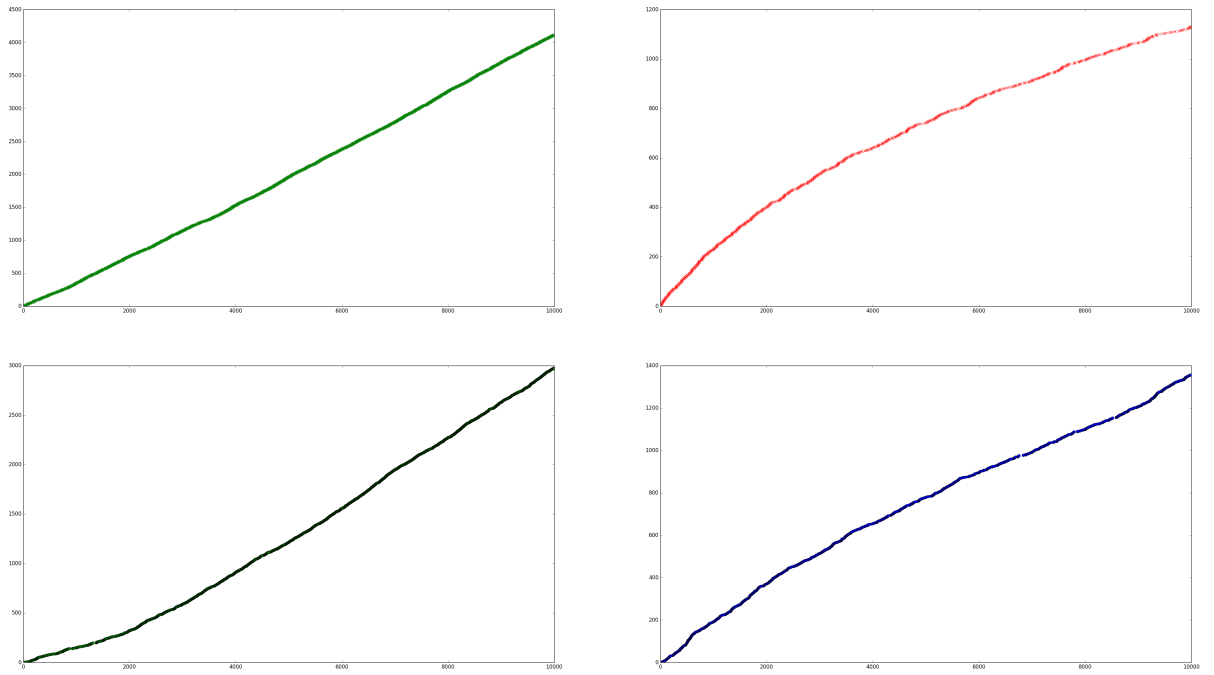


Figure 1: The plot of the training session of bot playing as 'X' in order of wins, loses, ties and invalid moves

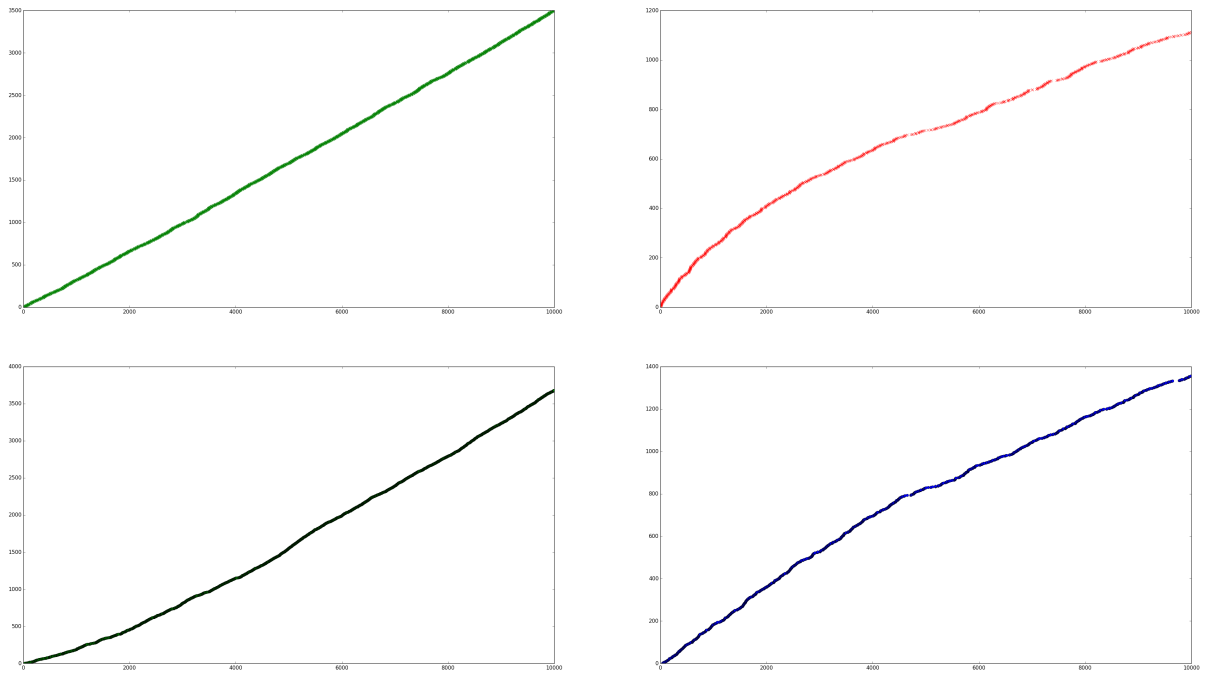


Figure 2: The plot of the training session of bot playing as 'O' in order of wins, loses, ties and invalid moves