

## Detailed Analysis of Assignment 2

By: Shravan

Player starts the game with first move:

Player: X

Compute: O

Terminal state: Win(+1), Lose(-1), draw(0)

- There are different cases of the play.
- The alphabet X and O is to represent whose move it is and the number leading to it, is to represent the order in which the moves were played.  
Ex: 1X means it is the first move by the player, 2O means first move by Computer.

Case 1(draw)			Case 2(draw)			Case 3(lose)		
9X	7X	4O	1X	7X	8O	5X	1X	4O
6O	1X	5X	4O	2O	5X		7X	3X
3X	8O	2O	3X	9X	6O	6O	2O	8O

Table : 1 Played using Min-Max

Case 1(draw)			Case 2(win)			Case 3(lose)		
1X	9X	6O	3X		4O	5X		3X
2O	3X	7X		5X			1X	
5X	8O	4O	2O		1X	2O	6O	4O

Table : 2 Played using Reinforcement Learning

### **Observation for Min-max(while Playing):**

- After playing the game using the min-max algorithm several times randomly, I noticed that after 2-3 moves by the player, the game reaches a state where the player knows that he cannot win the game. Either he can put some effort into making it a draw or the computer wins. The unbeatable nature of the game demonstrates that AI's decision-making, always selecting the best possible move to secure at least a draw.
- Computer will not lose with minmax, it either wins or it's a draw. Why? I analyzed from my data that minmax is always trying to minimize their losing cause and maximize its score, that is why we are reaching at most draw. For deciding the move, it explores through the values of all the possible states from the current state and chooses the minimum amongst all.
- The first move of the computer is fixed and has 2-3 positions which it gives preference to.

### **Observation for Reinforcement Learning(while playing):**

- Most of the time, the player wins the game if played intelligently, sometimes the match ends in a draw or a loss.
- Reaches saturation after a certain number of iterations.
- Computer can even win the game if the player does not play intelligently.
- I set the alpha value to 0.41 and analyzed that it becomes stable(achieves saturation) after number of terminals(958)\*200.
- When the player places its first and second move in the last and first cell of the board respectively then the player is more likely to win the game.
- The balance between exploration and exploitation helps to maintain the interest of the player so that some random moves help to reduce the best moves by computer.

### **Comparison:**

- Minmax explores the game tree to decide optimal moves whereas Reinforcement learning updates the values from the previous values and gets better with each iteration.
- Minmax guarantees an unbeatable strategy whereas the other one gradually approaches the unbeatable nature with certain no. of iterations.

- Minmax do not learn with iterations OR adapt, but reinforcement learns and adapts according to the environment.