ASSIGNMENT-2 REPORT

By: Kirtan Khichi

Introduction:

Tic Tac Toe is a simple two player game in which anyone wins when they have consecutive symbols (‘X’ or ‘O’) in vertical or horizontal or diagonally. It's a common thing that when two humans play the Tic Tac Toe game, they play it to pass their time and it becomes boring after some time, but the interesting thing takes place when a human can play the Tic Tac Toe with a computer and then humans also shows their interest in this, and which best move he/she can play to defeat the computer. For this task that computers play against humans we are using two algorithms MiniMax and Reinforcement learning.

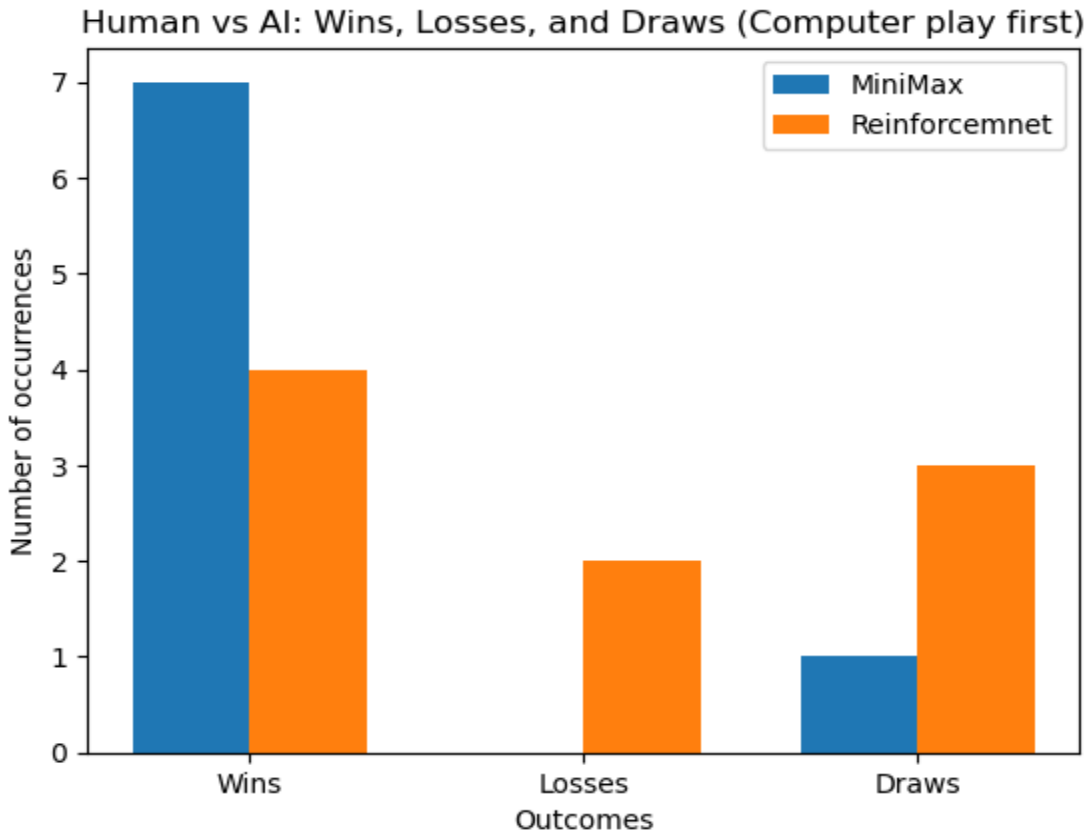
Minimax is a powerful tool for playing two-player games because it can help you make decisions that consider all possible outcomes.Minimax works by considering all possible moves you can make, then all possible moves your opponent can make, and so on, until the game ends. It then assigns a value to each possible outcome, such as +1 for winning, -1 for losing, and 0 for a draw. Finally, it chooses the move that leads to the best possible outcome, assuming that your opponent will also play optimally.

On the other hand, Reinforcement Learning is a powerful tool for teaching AI agents how to play games because it allows them to learn from their mistakes and improve over time. However, it can also be slow and require a lot of data to train.Reinforcement Learning works in a similar way. An AI agent is given a goal, such as winning a game of Tic Tac Toe. The agent then interacts with the environment, making decisions and observing the results. If the agent makes a good decision, it gets a reward. If it makes a bad decision, it gets a penalty. Over time, the agent learns to make better decisions in order to get more rewards.

Analysis of algorithms:

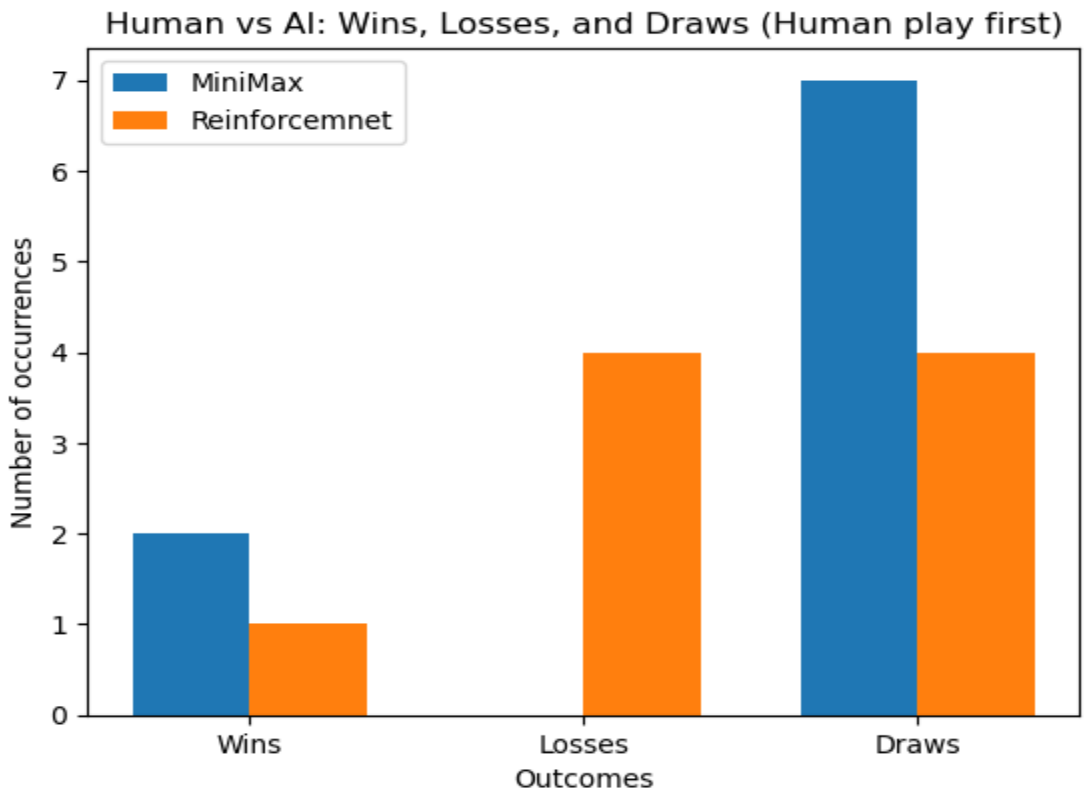
The bar graph shows the number of wins, losses, and draws in games of MiniMax and Reinforcement Learning (RL) against human players.

Case I: When the computer is playing first.



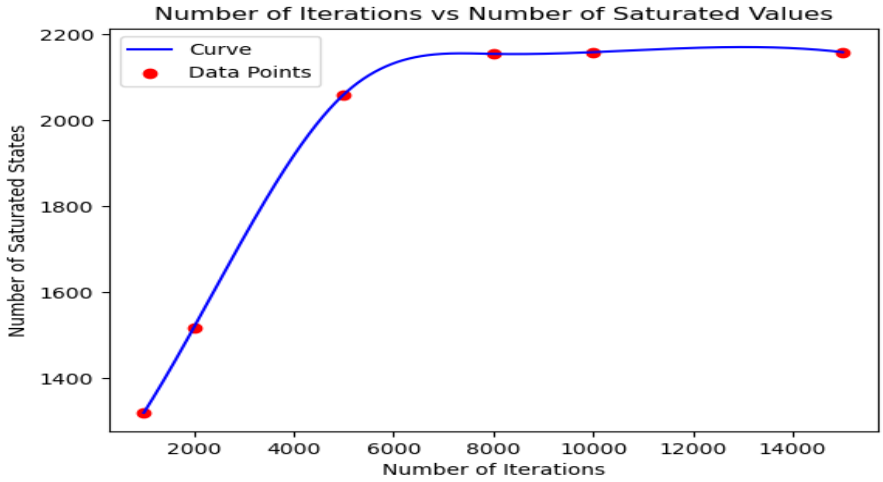
We can see in the above Histogram that when computer is starting the game to play first move by itself, then in this scenario for the case of MiniMax it will always play the game optimally and it will play the game with its best possible moves and will always wins the game or the game will be drawn and losses by MiniMax can be zero. When best move Human will play, He/She either draw the game or lose it .We can see this through the graph that MiniMax is either winning the game or the game is drawn, Whereas in the case of Reinforcement Learning computer is winning the game or the game is drawn , but here it is also losing the game against the humans because it is willing to take risks..We have analysed these case which we have shown it the above graph

Case II: **When Human is playing first.**



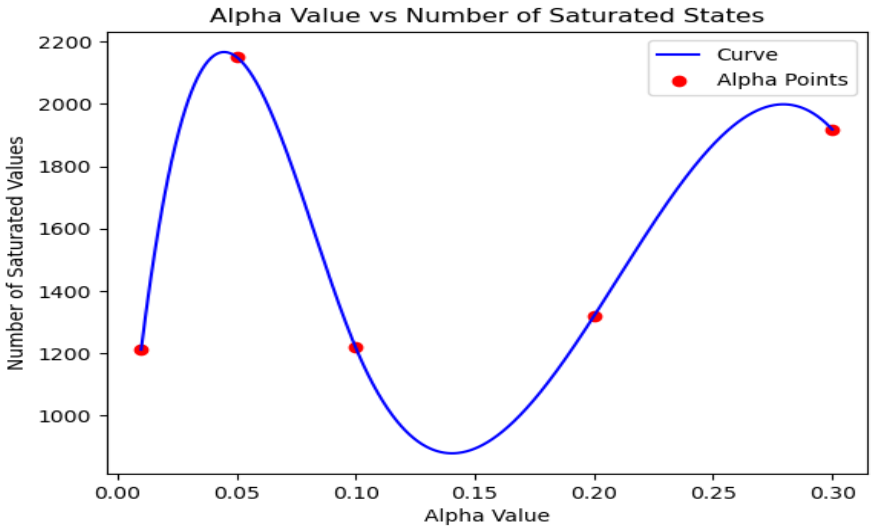
We can see in the above Histogram that when Human is starting the game to play first move by itself, then in this scenario for the case of MiniMax it will always play the game optimally and it will play the game with its best possible moves and will always wins the game or the game will be drawn and losses by MiniMax can be zero, But here the twist occurs that the probability of number of game drawn will increase. When best move Human will play, He/She either draw the game or lose it.We can see this through the graph that MiniMax is either winning the game or the game is drawn, Whereas in the case of Reinforcement Learning computer is winning the game or the game is drawn , but here it is also losing the game against the humans.We have analysed these case which we have shown it the above graph

Analysis of Saturation point in RL:



Keeping fix Alpha = 0.2 :

We can observe from the above curvature graph that when we are fixing the value of the alpha as 0.2 and increasing the number of iterations in the order of 1000,2000,5000,10000, 15000 respectively, then it's showing convergence, which means that we are achieving the saturation point in the training the Reinforcement learninig model.



Keeping fix Number of Iteration= 1000 :

We can observe from the above curvature graph that when we are fixing the value of the number of iteration as 1000 and increasing the alpha value in the order of 0.01,0.05,0.1,0.2, 0.3 respectively, then we can see that the best saturation point is on the alpha 0.05. From this we can infer that taking 0.05 is better choice for saturation point..

Iteration per state	MiniMax_play_first	RL_play_first	Remark
1000	MiniMax won	Minimax won	When we are doing more iteration per state, we are getting good value in the R.L.
3000	MiniMax won	MiniMax won	
80000	Minimax won	Draw	

Analysis of Json Storage:

Storage type	Memory Space	Remark
2-D array with nested parent-child relationship	208000 KB	Initially I was using 2D array as state and then I used 1-D array as state then I used String as states with nested parent-child relation. But, in the end I used only unique states-value pairs in the dictionary. From this I saved a lot of space.
1-D array with nested parent-child relationship	88000 KB	
States as string with parent-child relationship	38000 KB	
Unique states-value pair in dictionary	88 KB	