1. States: In the Tic-Tac-Toe game, nine cells will be in a 3 \* 3 board. Each cell reflects a unique arrangement of the board at any given point in time. The states will change after each user makes a decision. There will be 39 states in total.

Action: Action in the Tic-Tac-Toe game involves placing the agent’s marker ‘X’ or ‘O’ in one of the available empty spots. Actions are dynamically based on the current state of the board with legal moves. Actions can be represented either numerically with indices from 0 to 8 for the structure of a 3 \* 3 board or with a matrix system identified by row and column numbers. For example, the top-left corner is (0,0).

The actions are based on the current state. The state dictates which cells on the board are available for placing a marker. The agent must assess if the offensive moves to get to win state or do the defensive actions to block the opponent’s potential win. The action will update the board to reflect the new state.

1. Winning the game: Assign a reward of +1. Winning should be maximally encouraged because winning is the best outcome.

Losing the game: Assign a reward of -1 to lose the game. The loss should be discouraged because this leads to a loss.

Drawing the game: Assign a reward of +0.5 to the tie. Drawing a game is a neutral outcome and is preferable to a loss, so it cannot be -1.

Non-terminal moves: Assign a small negative number like -0.01, which is a small penalty. This setting is to encourage the agent to win with fewer moves, avoiding unnecessary moves.

This reward design will align the requirements of the game and encourage the agent to minimize the time of the game without changing the outcome.

1. I will change the value of the penalty in Non-terminal moves if the board is complex and larger. In our reward design, the small penalty was used to reduce unnecessary actions. We will increase the penalty to push the agent to the winning state in fewer moves. For example, changing the value to -0.03 or -0.05. This will help increase the pace and speed of playing the game. This approach will increase the efficiency and shorten the time.