1.1 Suppose, instead of playing against a fixed opponent, the reinforcement learning algorithm described above played against itself. What do you think would happen in this case? Would it learn a different way of playing?

SOLUTION: If the reinforcement learning algorithm described above played against itself instead of a fixed opponent in a game of tic-tac-toe, it would likely learn to play the game in a different way than if it were playing against a fixed opponent.

When playing against a fixed opponent, the algorithm is learning to optimize its strategy to beat that specific opponent. However, when playing against itself, the algorithm is essentially playing against multiple opponents with varying strategies, and it will need to learn to adapt to these different strategies.

As a result, the algorithm would likely develop a more generalized strategy for playing tic-tac-toe, rather than a strategy that is tailored to a specific opponent. This could result in a more balanced and robust playing style that is better suited to handle different opponents and game scenarios.

Additionally, playing against itself would allow the algorithm to continuously learn and improve its strategy over time, as it can always find new ways to challenge itself and test its limits. This could lead to the algorithm becoming an expert in playing tic-tac-toe, as it can learn from its mistakes and refine its strategy through self-play.