Suppose the reinforcement learning player was greedy, that is, it always played the move that brought it to the position that it rated the best. Would it learn to play better, or worse, than a nongreedy player? What problems might occur?

SOLUTION: If the reinforcement learning player was always greedy and chose the move that brought it to the position it rated the best, it would likely learn to play worse than a non-greedy player. The reason for this is that the greedy player would only consider immediate rewards and would not explore alternative moves that could lead to greater rewards in the long run.

A non-greedy player, on the other hand, would explore different moves and take risks that could lead to greater rewards in the long term. This would allow the non-greedy player to discover better strategies and avoid getting stuck in suboptimal moves that only provide short-term rewards.

The problem with the greedy approach is that it may result in the reinforcement learning player getting stuck in a local optimum and missing out on a better global optimum. For example, the greedy player may choose a move that provides an immediate reward, but that move could lead to a worse position in the future that reduces the overall reward. By exploring alternative moves, a non-greedy player could discover a better strategy that provides greater rewards in the long term.

In summary, while the greedy approach may provide some short- term gains, it is likely to lead to worse overall performance in the long term due to the player's lack of exploration and risk-taking.