Question:

Consider a simple grid world with the following properties:

- The grid is 4x4 (4 rows and 4 columns).
- The agent can take four possible actions in each state: up, down, left, or right.
- The goal is located at position (3, 3), and the agent receives a reward of 1 for reaching the goal and 0 for all other positions.
- The agent's objective is to find the optimal policy that maximizes long-term rewards using Value Iteration.

Given this scenario:

1. What is the purpose of the get_next_state function in the grid world?
2. What is the significance of the discount factor (gamma) set to 0.9 in this scenario?
3. Explain how the Value Iteration process works and how it contributes to finding the optimal policy in this grid world example.
4. What is the resulting value function (V), and what does it represent?
5. Based on the final output of the optimal policy, how should the agent navigate the grid to reach the goal?