Exercise 4.10

$$V_{k+1}(s) = \max_a \sum_{s',r} p(s',r \mid s,a) \cdot [r + \gamma V_k(s')]$$

Q-value iteration update, which updates the Q-values for each state-action pair. The update equation for Q-value can be derived from Bellman equation.

$$Q_{k+1}(s,a) = \sum_{s',r} p(s',r \mid s,a) \cdot [r + \gamma \cdot \max_{a'} Q_k(s',a')]$$

Analog of the value iteration update for action values would be

$$q_{k+1}(s,a) = \sum_{s',r} p(s',r \mid s,a) \cdot [r + \gamma \cdot \max_{a'} q_k(s',a')]$$

Q-values, $q_{k+1}(s,a)$ represent the Q-value for taking action in state s at iteration $k+1$.