COMPSCI 687 Pop Quiz 5

Name: _____

Instructions. This quiz is closed book and closed notes. You may not communicate with any other students or use any electronic devices. You have **5 minutes** to complete the quiz. As discussed in class, answers must be **perfect down to the symbol** to receive credit. Each question is graded as either correct or incorrect; there is no partial credit. **If your name is not legible, you will receive a 0 on the quiz.** You may omit quantifiers and the sets in operator subscripts (e.g., writing $\sum_a$ instead of $\sum_{a \in \mathcal{A}}$), but the indices themselves are required.

1. Give the Bellman equation for v^π as a recursive expression for v^π using R, p, π, and γ.

   $v^\pi(s) =$

2. Give the Bellman optimality equation for v^* as a recursive expression for v^* using R, p, and γ.

   $v^*(s) =$

3. Give the dynamic programming policy evaluation update for a fixed policy π, using R, p, π, γ, and the current estimate v_k.

   $v_{k+1}(s) =$

4. In policy iteration, the current policy is π_k. Give the expression used to obtain the next deterministic policy greedily from v^{π_k}, using R, p, and γ.

   $\pi_{k+1}(s) \in$