

Name _____

CMPSCI 687 Pop Quiz 6

Instructions: This quiz is **closed** notes—do not use your notes or a laptop. Do not discuss problems with your neighbors until after everyone has handed in their quiz.

Note: Your answers should use math (equations), not English.

1. The temporal difference (TD) error, given that the agent took action a in state s , resulting in it getting reward r and transitioning to state s' is defined as:

$$\delta =$$

2. The temporal difference learning update given that the agent took action a in state s , resulting in it getting reward r and transitioning to state s' is defined as:

$$v(s) \leftarrow$$

3. A positive TD-error means that:
 - a) The actual outcome appears to be better than what was predicted.
 - b) The actual outcome appears to be worse than what was predicted.
4. (True or False) If we do not know P and R , but we know v^* , then we can compute an optimal action, a , for any state, s .