

Name _____

CMPSCI 687 Pop Quiz 7

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. In class we used a property of natural logarithms (which we will use again in the future). That property was:

$$\frac{\partial \ln(f(x))}{\partial x} =$$

2. (True or False) All actor-critic algorithms are policy gradient algorithms
3. (True or False) All policy gradient algorithms are actor-critic algorithms.
4. (True or False) Actor-critic is a specific algorithm, which was written on the board on class, not a class of algorithms.
5. The actor in an actor-critic:
 - a) Criticizes the decisions made by the critic.
 - b) Implements a policy—it selects actions when given a state.
 - c) Computes the temporal-difference error given a tuple (s, a, r, s') .
 - d) Estimates the state-value function.
6. The critic in an actor-critic:
 - a) Criticizes the decisions made by the actor.
 - b) Implements a policy—it selects actions when given a state.
 - c) Computes $\frac{\partial \ln(\pi(s, a, \theta))}{\partial \theta}$.
 - d) Estimates the state-value function.
7. Write the definition of the temporal difference error below (using the state-value function, not the action-value function).