Name _____

CMPSCI 687 Pop Quiz 2

Instructions: You have 10 minutes to complete this quiz. 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.

For the first two problems, circle the **single** letter corresponding to the most appropriate answer. For the third problem, circle true or false to indicate whether the statement is true or false.

1. In *black-box optimization* (BBO) for policy search, the term "black-box" refers to:
   (a) The fact that the policy is a black box—we do not know how changes to the policy parameter vector, θ, will change the distribution over actions, and thus the policy is a black box.
   (b) The fact that we view the agent as a black box—we can sample actions from its current policy, but we do not know the agent's exact policy (the distribution over actions that it is currently using for each state).
   (c) The fact that we treat the objective function, J, as a black-box that we do not know the inner workings of. For example, we do not have an analytic expression for $J(\pi)$.
   (d) The fact that BBO algorithms were invented by Leemon Blackbox.
   (e) The fact that we view the environment as a black box—we do not know the transition function, P, or reward function, R. Instead we can only interact with the environment as a "black-box" that we do not know the internal workings of.
   (f) All of the above.

2. A state representation is Markovian if:
   (a) For all s, a, s', t, and i:
   $$\Pr(S_{t+1} = s' | S_t = s, A_t = a) = \Pr(S_{i+1} = s' | S_i = s, A_i = a).$$
   (b) For all s, a, s', and h:
   $$\Pr(S_{t+1} = s' | H_{t-1} = h, S_t = s, A_t = a) = \Pr(S_{t+1} = s' | S_t = s, A_t = a).$$
   (c) For all s and a:
   $$\sum_{s' \in \mathcal{S}} P(s, a, s') = 1.$$
   (d) For all s, a, and h:
   $$\Pr(A_t = a | H_{t-1} = h, S_t = s) = \Pr(A_t = a | S_t = s)$$

3. (True or False) Every finite MDP with bounded rewards and $\gamma < 1$ has one, and only one, optimal policy.