Name _____

CMPSCI 687 Pop Quiz 3

Instructions: You have 5 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. The Markov property states that

   (a) The present is independent of the past given the future.

   (b) The future is independent of the past given the present.

   (c) The future is independent of the present given the past.

   (d) The past is independent of the present given the future.

   (e) All of the above.

2. The state-value function is defined as:

   (a) $v^\pi(s) = \mathbf{E}\left[\sum_a \pi(s,a) \sum_{s'} P(s,a,s')R(s,a,s')\right]$.

   (b) $v^\pi(s) = \mathbf{E}\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k} \middle| S_t = s, \pi\right]$.

   (c) $v^\pi(s) = \sum_a \pi(s,a) \sum_{s'} P(s,a,s')R(s,a,s')$.

   (d) $v^\pi(s,a) = \mathbf{E}\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k} \middle| S_t = s, A_t = a, \pi\right]$, where conditioning on π means that all actions *after* A_t are sampled according to π. If the agent returns to state s at a time $i > t$, then it will apply policy π.

   (e) $v^\pi(s,a) = \mathbf{E}\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k} \middle| S_t = s, A_t = a, \pi\right]$, where conditioning on π means that all actions *after* A_t are sampled according to π. If the agent returns to state s at a time $i > t$, then it will take action a regardless of whether π would select action a in state s.

3. (True or False) The value of a state can depend on the initial state distribution.