hw4_mdps_q4_value_iteration_properties

Question 4: Value Iteration Properties

0.0/5.0 points (graded)

Which of the following are true about value iteration? We assume the MDP has a finite number of actions and states, and that the discount factor satisfies $0 < \gamma < 1$.

- [x] Value iteration is guaranteed to converge. ✔
- [x] Value iteration will converge to the same vector of values (V^*) no matter what values we use to initialize V. ✔
- [ ] None of the above

1. For discount less than 1, value iteration is guaranteed to converge.
2. At convergence, the following equation must be satisfied for all states:

$V^*(s) = max_a \sum_{s'} T(s,a,s') [R(s,a,s') + \gamma V^*(s')]$

There will only be one set of values that satisfies this condition, so no matter where we start value iteration, we will always arrive at the same set of values on convergence.

Submit

Answers are displayed within the problem