
Algorithm 1 Value Iteration

- 1: For each state s , initialize $V(s) := 0$.
- 2: **for** until convergence **do**
- 3: For every state, update

$$V(s) := R(s) + \max_{a \in A} \gamma \sum_{s'} P_{sa}(s') V(s').$$
