

```

1:   Algorithm MDP_value_iteration():
2:     for all  $x$  do
3:        $\hat{V}(x) = r_{\min}$ 
4:     endfor
5:     repeat until convergence
6:       for all  $x$ 
7:         
$$\hat{V}(x) = \gamma \max_u \left[ r(x, u) + \int \hat{V}(x') p(x' | u, x) dx' \right]$$

8:       endfor
9:     endrepeat
10:    return  $\hat{V}$ 

```

```

1:   Algorithm MDP_discrete_value_iteration():
2:     for  $i = 1$  to  $N$  do
3:        $\hat{V}(x_i) = r_{\min}$ 
4:     endfor
5:     repeat until convergence
6:       for  $i = 1$  to  $N$  do
7:         
$$\hat{V}(x_i) = \gamma \max_u \left[ r(x_i, u) + \sum_{j=1}^N \hat{V}(x_j) p(x_j | u, x_i) \right]$$

8:       endfor
9:     endrepeat
10:    return  $\hat{V}$ 

```

```

1:   Algorithm policy_MDP( $x, \hat{V}$ ):
2:     return 
$$\operatorname{argmax}_u \left[ r(x, u) + \sum_{j=1}^N \hat{V}(x_j) p(x_j | u, x) \right]$$


```

**Table 14.1** The value iteration algorithm for MDPs, stated here in its most general form and for MDPs with finite state and control spaces. The bottom algorithm computes the best control action.