Algorithm MDP_value_iteration( ):

for all x *do*
  $\hat{V}(x) = r_{\min}$
endfor
repeat until convergence
  for all x
    $$\hat{V}(x) = \gamma \max_u \left[r(x,u) + \int \hat{V}(x')\, p(x' \mid u, x)\, dx' \right]$$
  endfor
endrepeat
return $\hat{V}$

Algorithm MDP_discrete_value_iteration( ):

for $i = 1$ *to* N *do*
  $\hat{V}(x_i) = r_{\min}$
endfor
repeat until convergence
  for $i = 1$ *to* N *do*
    $$\hat{V}(x_i) = \gamma \max_u \left[r(x_i,u) + \sum_{j=1}^{N} \hat{V}(x_j)\, p(x_j \mid u, x_i) \right]$$
  endfor
endrepeat
return $\hat{V}$

Algorithm policy_MDP($x, \hat{V}$):

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

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.