

MDP solutions

February 2019

Value Iteration for Grid (slide 7)

$$\begin{aligned} Q_3((3,3), right) &= \sum_s^I T((3,3), right, s') [R((3,3), right, s') + \gamma V_2(s')] \\ &= 0.8 * (0 + .9 * 1) + 0.1(0 + 0.9 * 0.72) + 0.1(0 + 0.9 * 0) \end{aligned}$$

Value Iteration for Car (slide 12)

$$\begin{aligned} Q_{k+1}(cool, fast) &= T(cool, fast, cool) [R(cool, fast, cool) + \gamma * V_k(cool)] + \\ &T(cool, fast, warm) [R(cool, fast, warm) + \gamma * V_k(warm)] \\ Q_1(cool, fast) &= 0.5[2 + 0] + 0.5[2 + 0] = 2, v_1(cool) = 2 \end{aligned}$$

$$\begin{aligned} Q_{k+1}(warm, slow) &= T(warm, slow, cool) [R(warm, slow, cool) + \gamma * V_k(cool)] \\ &+ T(warm, fast, o.h.) [R(warm, slow, warm) + \gamma * V_k(warm)] \\ Q_1(warm, slow) &= 0.5[1 + 0] + 0.5[1 + 0] = 1, V_1(warm) = 1 \end{aligned}$$

$$\begin{aligned} Q_2(cool, fast) &= 0.5[2 + 1 * 2] + 0.5[2 + 1.1] = 3.5, V_2(cool) = 3.5 \\ Q_2(warm, slow) &= 0.5[1 + 1 * 2] + 0.5[1 + 1 * 1] = 2.5, V_2(warm) = 2.5 \end{aligned}$$