

Starting w

$$q_w = q_\pi$$

$$\pi = \varepsilon\text{-greedy}(q_w)$$

$$q_w \approx q^*$$

