

Class 9

Shikhar Saxena

February 07, 2023

Contents

Infinite Horizon Problem	1
MDP- Expected Discounted Cost.....	1

Infinite Horizon Problem

Q learning Algorithm:

$$Q(S_t, A_t) \rightarrow Q(S_t, A_t) + \alpha[R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t)]$$

(Empirical mean update formula sort of used here)

MDP- Expected Discounted Cost

- Assume Rewards are bounded.