

Decision-time Planning

Model Predictive Control (MPC)

$$\pi_{\text{MPC}}(s; f, r, Q_{\theta}) = \arg \max_{a_0} \max_{a_1, \dots, a_{H-1}} = \sum_{t=0}^{H-1} \gamma^t r(s_t, a_t) + \gamma^H Q_{\theta}(s_H, a_H) \quad \text{s.t.} \quad s_{t+1} = f(s_t, a_t)$$
$$s_0 = s$$

Decision-time Planning

Main Takeaways