

$$\max_{\theta \in \mathbb{R}^d} \mathbb{E}_{a_t \sim \pi_{\theta}(\cdot | s_t)} \left[\sum_{t=0}^{T-1} \gamma^t r(s_t, a_t) \mid s_0 = s \right],$$