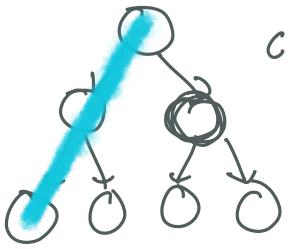


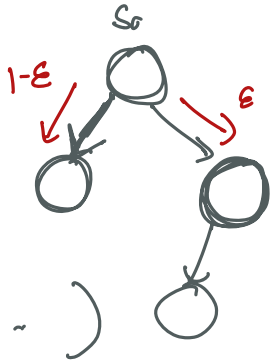
$$c(s,a) = \mathbb{1}(a \neq \pi^*(s))$$

Assume that π with probability ε takes the wrong action.
prob $1-\varepsilon$, takes the right action
ON EXPERT STATES

$$\boxed{J(\pi) - J(\pi^*)}$$

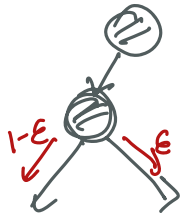
$$J(\pi^*) = \sum_{t=0}^{T-1} \underset{s_t \sim d_t^{\pi^*}}{E} \underbrace{c(s_t, \pi^*(s_t))}_{0} = 0$$

$$J(\pi) = \sum_{t=0}^{T-1} \underset{s_t \sim d_t^{\pi}}{E} c(s_t, \pi(s_t))$$



$$= \varepsilon \Big(\underbrace{1 + 1 + 1 + 1 + 1 \cdots}_{T} \Big)$$

$$+ (1-\varepsilon)\Big(0 + \varepsilon \underbrace{(1 + 1 + 1 + \cdots\cdots)}_{T-1}$$



$$+ (1-\varepsilon)(0 + \cdots\cdots\cdots)$$

$$= \varepsilon T + (1-\varepsilon)\varepsilon(T-1) + (1-\varepsilon)^2 \varepsilon (T-2) + \cdots\cdots$$

$$= \varepsilon \Big[T + \underbrace{(1-\varepsilon)}_{\leq 1}(T-1) + \underbrace{(1-\varepsilon)^2}_{\leq 1}(T-2) + \cdots\cdots \Big]$$

$$\leq \varepsilon \big[T + (T-1) + (T-2) + \cdots\cdots 1 \big]$$

$$\leq \varepsilon \frac{T(T+1)}{2} = \varepsilon \frac{(T^2+T)}{2} \approx O(\varepsilon T^2)$$