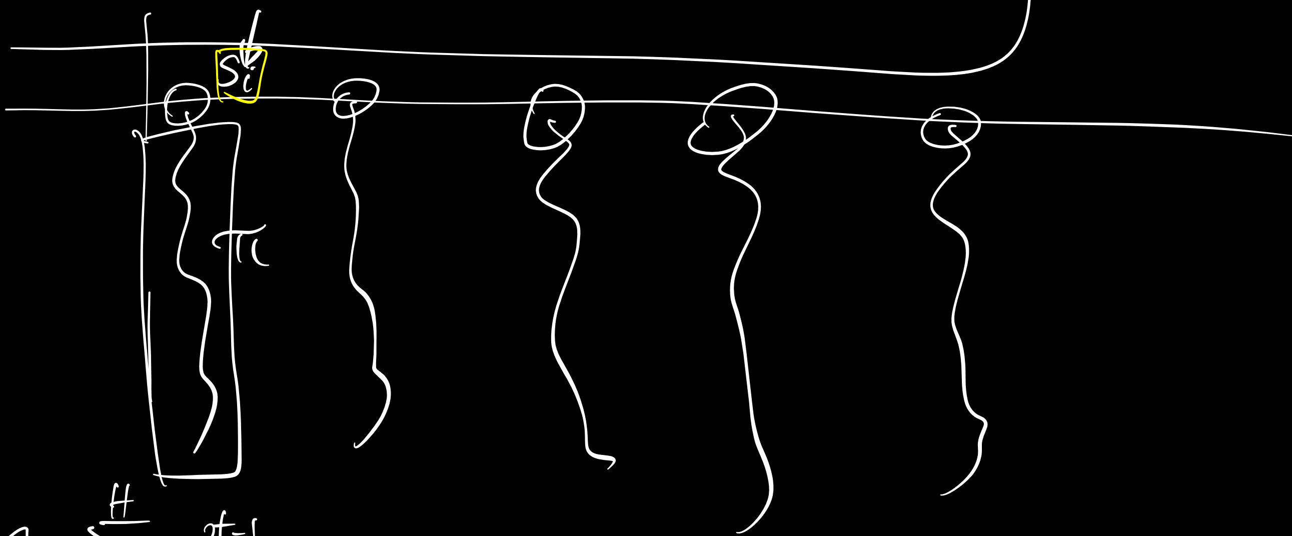
Learn V^π for given π.

Maintain $V \in \mathbb{R}^{|S|}$.



$G_i = \sum_{t=1}^{H} \gamma^{t-1} r_t$.

$$V(s_i) \leftarrow V(s_i) + \alpha (G_i - V(s_i)).$$

$$V(s) \longrightarrow \mathbb{E}[G \mid s] = V^\pi(s).$$

$$V(s_t) \leftarrow V(s_t) + \alpha (G_t - V(s_t))$$

$s_1, a_1, r_1 \cdots s_t, a_t, r_t, s_{t+1}, \ldots$

$$G_t = r_t + \gamma V(s_{t+1}).$$

$$\mathbb{E}[G_t \mid s_t] = (\mathcal{T}^\pi V)(s_t)$$

$$= R(s,\pi) + \gamma \mathbb{E}_{s' \sim P(s,\pi)}[V(s')].$$

H

s_t

2-step TD

$$G_t = r_t + \gamma r_{t+1} + \gamma^2 V(s_{t+2}).$$

$$\mathbb{E}[G_t \mid s_t = s] = \mathbb{E}[r_t + \gamma r_{t+1} + \gamma^2 V(s_{t+2}) \mid s_t = s].$$

$$= R(s, \pi) + \gamma \mathbb{E}[r_{t+1} + \gamma V(s_{t+2}) \mid s_t = s].$$

$$= R(s,\pi) + \gamma \mathbb{E}[\mathbb{E}[r_{t+1} + \gamma V(s_{t+2}) \mid s_t = s, s_{t+1}] \mid s_t = s].$$

$$= R(s,\pi) + \gamma \mathbb{E}[\underline{\mathbb{E}[r_{t+1} + \gamma V(s_{t+2}) \mid s_{t+1}]} \mid \underline{s_t = s}]$$

$$= R(s,\pi) + \gamma \mathbb{E}[(T^\pi V)(s_{t+1}) \mid \underline{s_t = s}].$$

$$= \underbrace{R(s,\pi) + \gamma \mathbb{E}_{s' \sim P(\cdot \mid s, \pi)}[\boxed{(T^\pi V)}(s')]}_{f}.$$

$$= (T^\pi f)(s) = (T^\pi(T^\pi V))(s)$$

$$= (T^{\pi})^2 V(s).$$