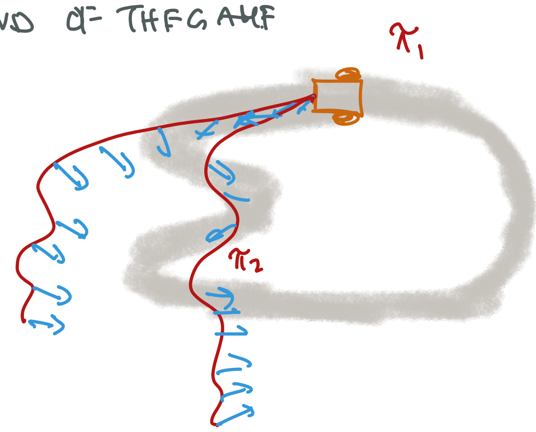
Imitation Learning - Interactive

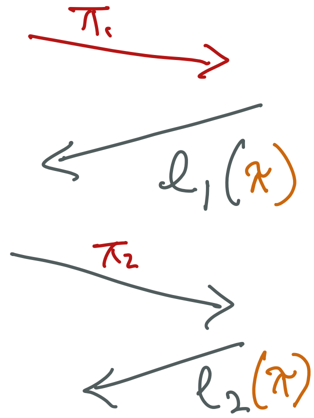
Define the loss function at every round of the game



Dataset 1:

$$\{ s_1, a_1^*, s_2, a_2^*, \ldots \}$$

$$\ell_1(\pi) = \sum_{t=0}^{T-1} \mathbb{E}_{s_t \sim d_t^{\pi_1}} \mathbb{1}\left(\pi(s_t) \neq a_t^*\right)$$



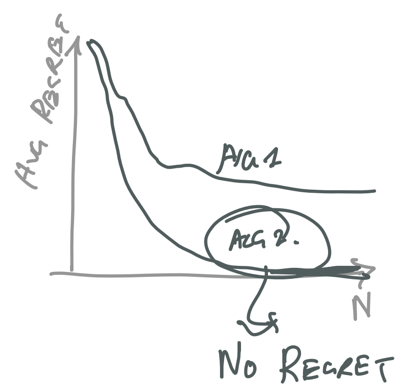
Dataset 2:

$$\{ s_1, a_1^*, s_2, a_2^*, \ldots \}$$

$$\ell_2(\pi) = \sum_{t=0}^{T-1} \mathbb{E}_{s_t \sim d_t^{\pi_2}} \mathbb{1}\left(\pi(s_t) \neq a_t^*\right)$$

Policies (Π)

	Round 1	Round 2	Round 3	...	Round N
π^1	0.2	0.7	1		
π^2	0.1	1.0	1		
π^3	1.0	0.3			
⋮	⋮	⋮	0.2		
⋮	⋮	⋮	0.3		
π^k	⋮	⋮	1		



$\pi_1 = \pi^3$

$$\text{Average Regret} = \frac{1}{N}\left(\sum_{i=1}^{N} \ell_i(\pi_i) - \min_{\pi \in \Pi} \sum_{i=1}^{N} \ell_i(\pi) \right)$$

$1 + 1 + 0.3 + \cdots$

$$\text{Best Response:} \quad \pi_i = \underset{\pi \in \Pi}{\arg\min}\ \ell_{i-1}(\pi)$$

FOLLOW THE LEADER (FTL) $\quad \pi_i = \underset{\pi \in \Pi}{\operatorname{argmin}} \sum_{j=1}^{i-1} l_j(\pi)$

Is DAGGER FTL?

DAGGER

Collect D_i with π_i

$D \leftarrow D \cup D_i$

$\pi_{i+1} \leftarrow \text{Train}(D)$

$$\pi_{i+1} = \underset{\pi \in \Pi}{\arg\min} \sum_{j=1}^{i} \sum_{t=0}^{T-1} \underset{s_t \sim D_i}{E} \mathbb{1}\left(\pi(s_t) \neq a_t^*\right)$$

$\downarrow$

$s_t \sim d_t^{\pi_i}$

$$\boxed{l_i(\pi) = \sum_{t=0}^{T-1} \underset{s_t \sim d_t^{\pi_i}}{E} \mathbb{1}\left(\pi(s_t) \neq a_t^*\right)} = \underset{\pi \in \Pi}{\arg\min} \sum_{j=1}^{i} l_j(\pi)$$

DAGGER RETURNS A SERIES OF POLICIES $\pi_1, \pi_2 \ldots \pi_N$

$\exists$ a policy π_i, $l_i(\pi_i)$ is SMALL?

$$\min_{i=1,\ldots N} l_i(\pi_i) \leq \frac{1}{N}\sum_{i=1}^{N} l_i(\pi_i)$$

$$\leq \boxed{\frac{1}{N}\sum_{i=1}^{N} l_i(\pi_i) - \frac{1}{N}\min_{\pi \in \Pi}\sum_{i=1}^{N} l_i(\pi)}$$

$$+ \frac{1}{N}\min_{\pi \in \Pi}\sum_{i=1}^{N} l_i(\pi^*)$$

$$\leq \text{AVG REGRET} + \frac{1}{N} \min_{\pi \in \Pi} \sum_{i=1}^{N} l_i(\pi^*)$$

$\longrightarrow 0$ as $N \to \infty$

$\preceq O(\epsilon H T)$