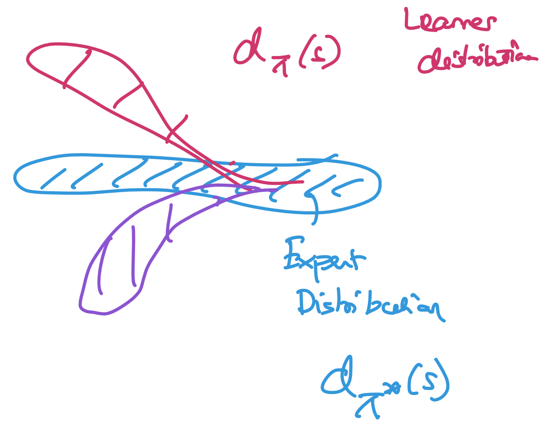
CHICKEN-OR-EGG

LEARNER DISTRIBUTION



ALGORITHM

1. Start with some policy π
2. Rollout policy π
3. See state it visited $\sim d_\pi(s)$ $\rightsquigarrow$ QUERY EXPERT $a^* \leftarrow \pi^*(s)$
4. Train on new data.

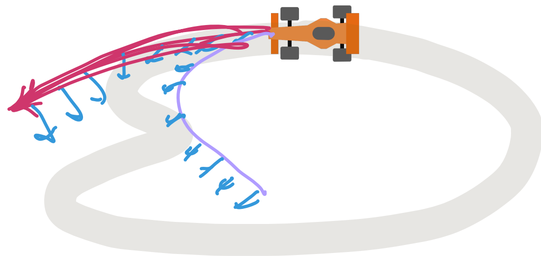
PROPOSED ALGORITHM

INIT: RANDOM POLICY π_0

for $i = 0 \ldots (N-1)$

- Collect data of state π_i visits

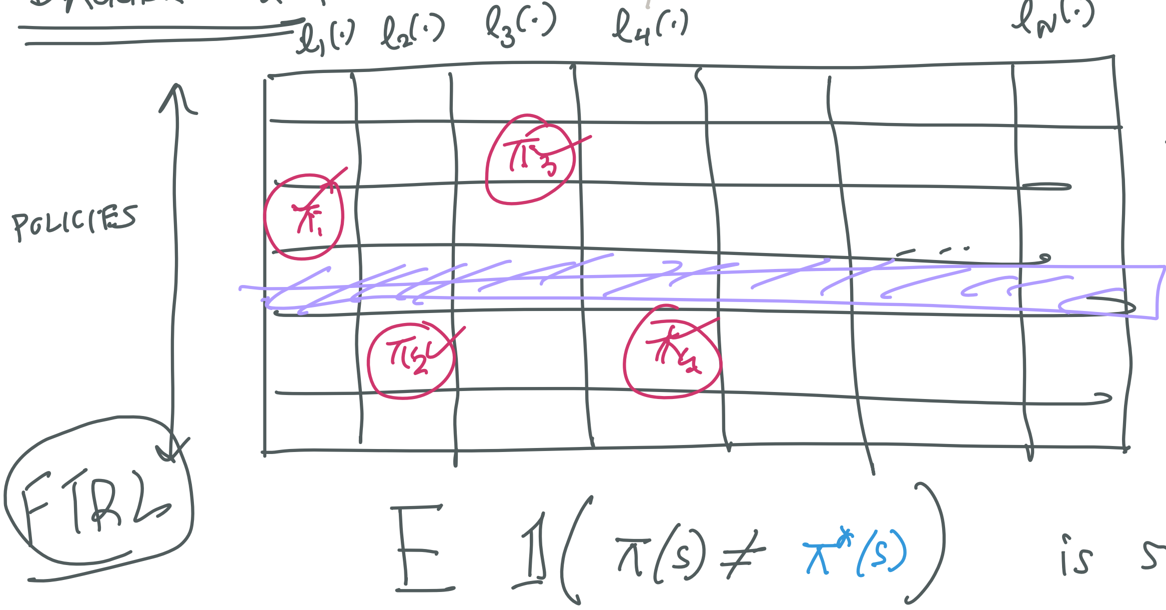
  $s \sim d_{\pi_i}(s)$



- Query interactive expert to get $\pi^*(s) \rightarrow a^*$

- TRAIN π_{i+1} on (s, a^*)

DAGGER PROOF



$$\ell_i(\pi) = \mathop{\mathbb{E}}_{s \sim d_{\pi_i}} \mathbb{1}\left(\pi(s) \neq \pi^*(s)\right)$$

$$\mathop{\mathbb{E}}_{s \sim d_{\pi}} \mathbb{1}\left(\pi(s) \neq \pi^*(s)\right) \text{ is small}$$

Is there a round i where $\underline{\ell_i(\pi_i)}$ is small.

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

$$\leq \underline{\left[\frac{1}{N} \sum_{i=1}^{N} \ell_i(\pi_i) - \frac{1}{N} \min_{\pi^* \in \Pi} \sum_{i=1}^{N} \ell_i(\pi^*)\right]} + \frac{1}{N} \min_{\pi^* \in \Pi} \sum_{i=1}^{N} \ell_i(\pi^*)$$

$$\leq \frac{1}{N} \mathrm{REG}(\cdot) + \boxed{\frac{1}{N} \min_{\pi^* \in \Pi} \sum_{i=1}^{N} \ell_i(\pi^*)}$$

$$\lim_{N \to \infty} \frac{1}{N} \mathrm{REG}(\cdot) \to 0 \qquad\qquad \leq \varepsilon$$