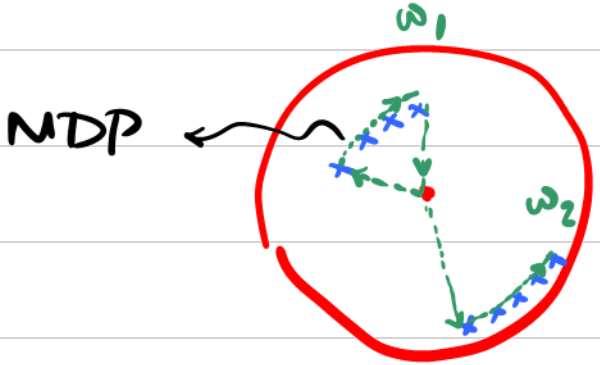
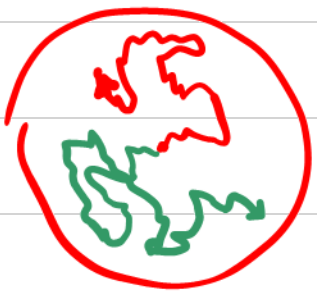
Meta RL





* I don't want to solve MDPs in separation.
* I want to devise a strategy to solve them together.

$$M_1, M_2, \dots, M_k \sim p(M) \quad \pi_i = \underbrace{f_\theta}_{\text{RL Alg. to solve the } i\text{th task}}(M_i) \qquad \max_\theta \frac{1}{k} \sum_{i=1}^{k} J_{\pi_i}(\theta)$$

$M =$

$$\log p(\text{target} \mid s_t, a_t, s_{t+1}) - \log p(\text{target} \mid s_t, a_t)$$

$$= \log \frac{p(\text{target} \mid s_t, a_t, s_{t+1})}{p(\text{target} \mid s_t, a_t)}$$

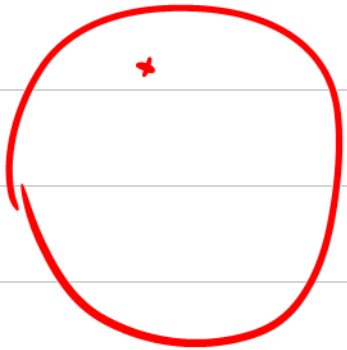
$$p(\text{target} \mid s_t, a_t, s_{t+1}) = \frac{p(s_t, a_t, s_{t+1} \mid \text{target}) \cdot p(\text{target})}{p(s_t, a_t, s_{t+1})}$$

$$p(\text{target} \mid s_t, a_t) = \frac{p(s_t, a_t \mid \text{target}) \cdot p(\text{target})}{p(s_t, a_t)}$$

$$p(s_{t+1} \mid s_t, a_t)$$

$$p_{\text{target}}(s_{t+1} \mid s_t, a_t)$$

Scenario 1.



Scenario 2.

Scenario 3.