

$$\nabla_{\theta_{k+1}^Q} \mathcal{L}(\theta_{k+1}^Q) = \mathbb{E}_{(s,a,r,s') \sim \mathcal{U}(D)} \left[\underbrace{(Q(s,a|\theta_{k+1}^Q) - y)}_{\text{regret}} \underbrace{\nabla_{\theta_{k+1}^Q} Q(s,a|\theta_{k+1}^Q)}_{\text{averseness}} \right]$$

parameter update (backprop!)

