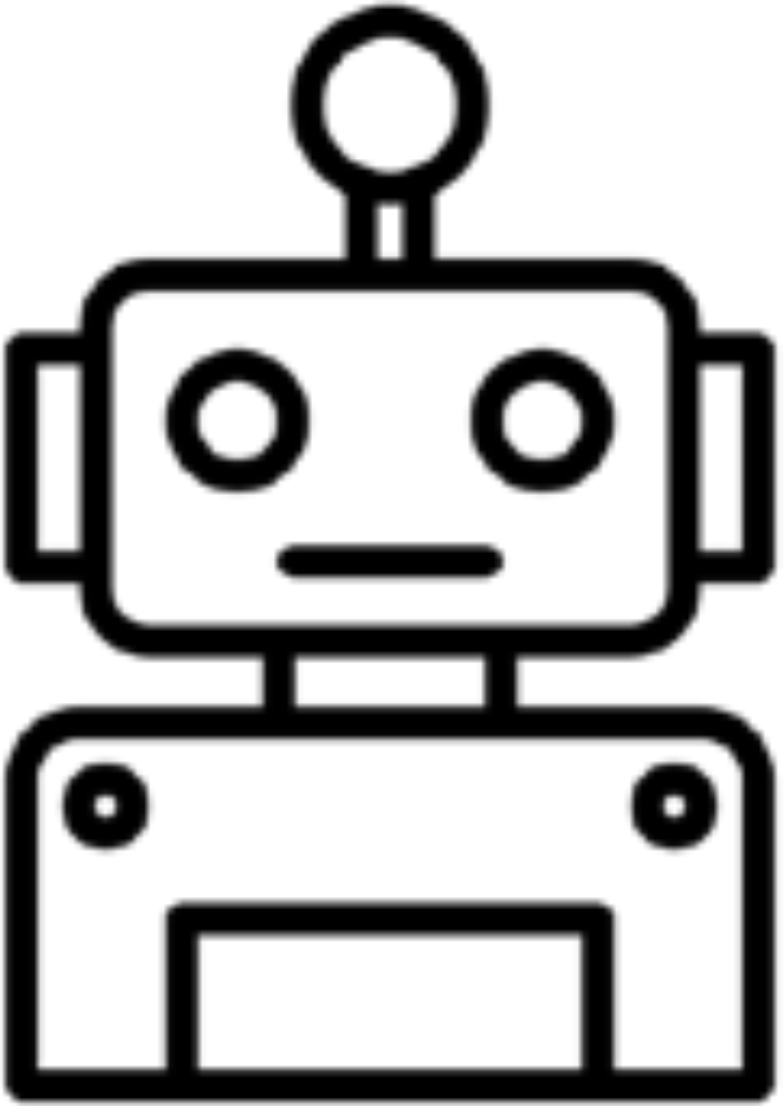
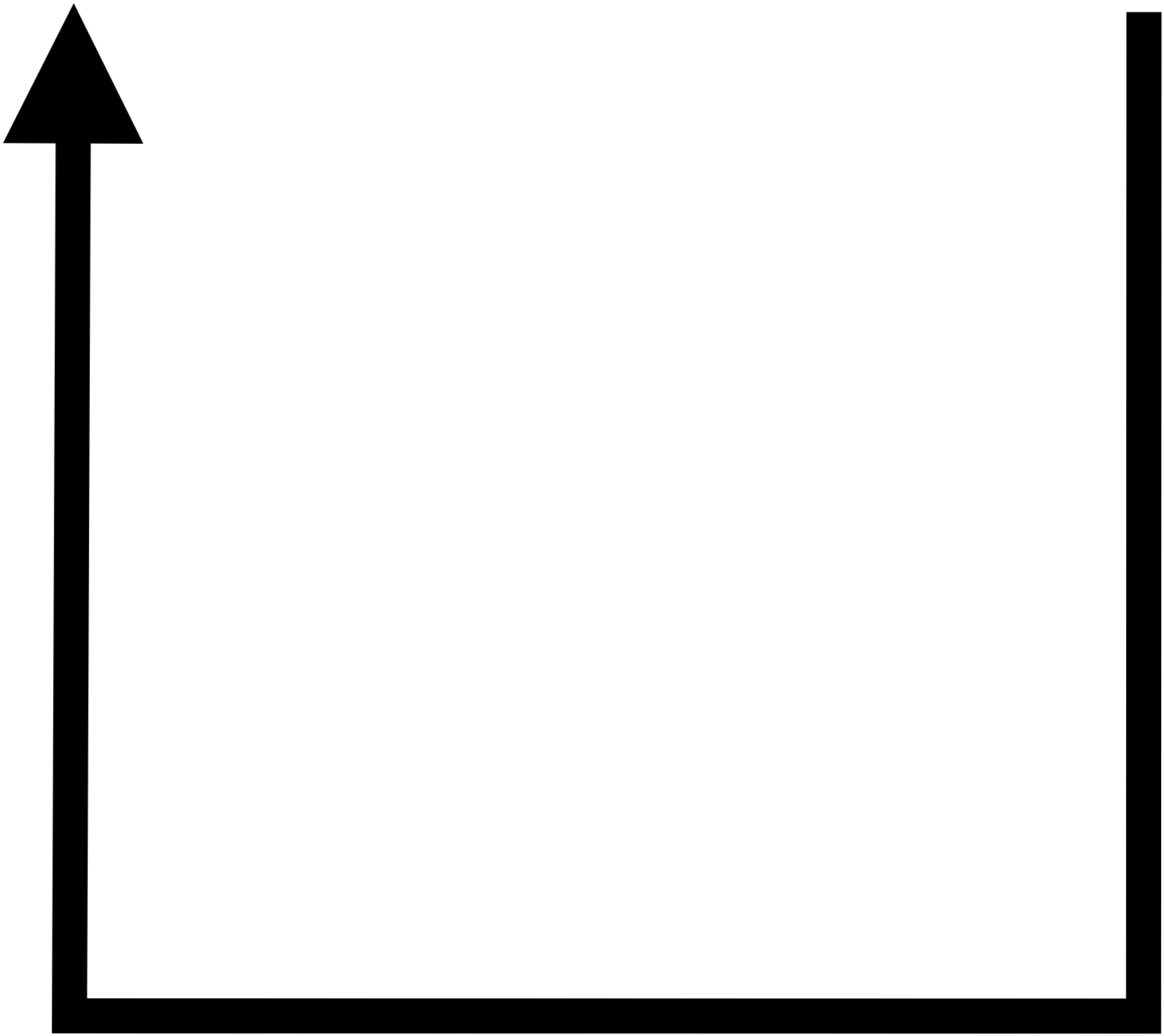
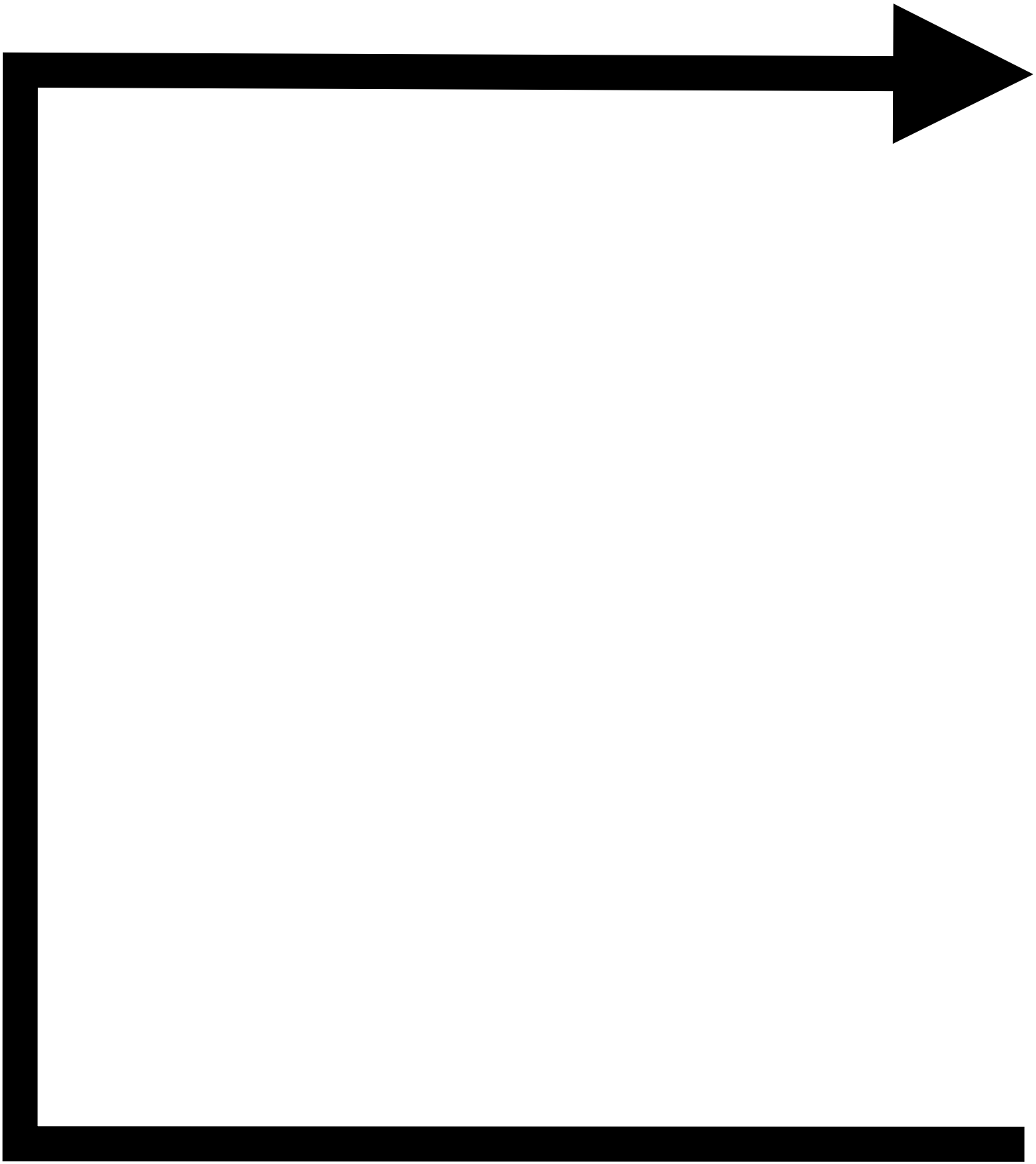
Reinforcement Learning

7







$$a = \pi(s)$$

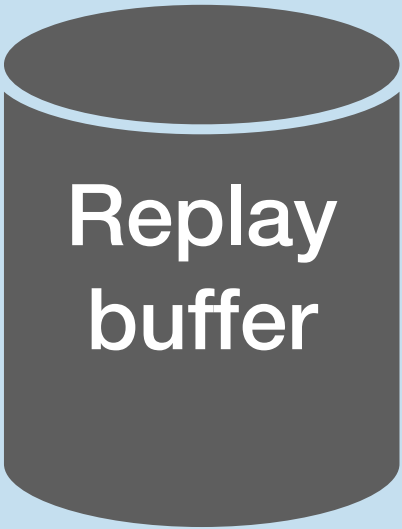
Actions

$$s' \sim P(\cdot \mid s, a)$$

Transition function

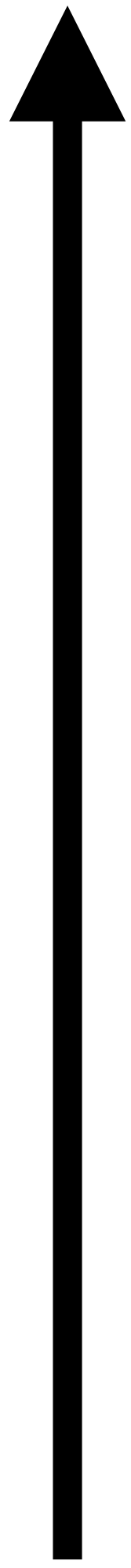
$s', \quad r(s,a)$

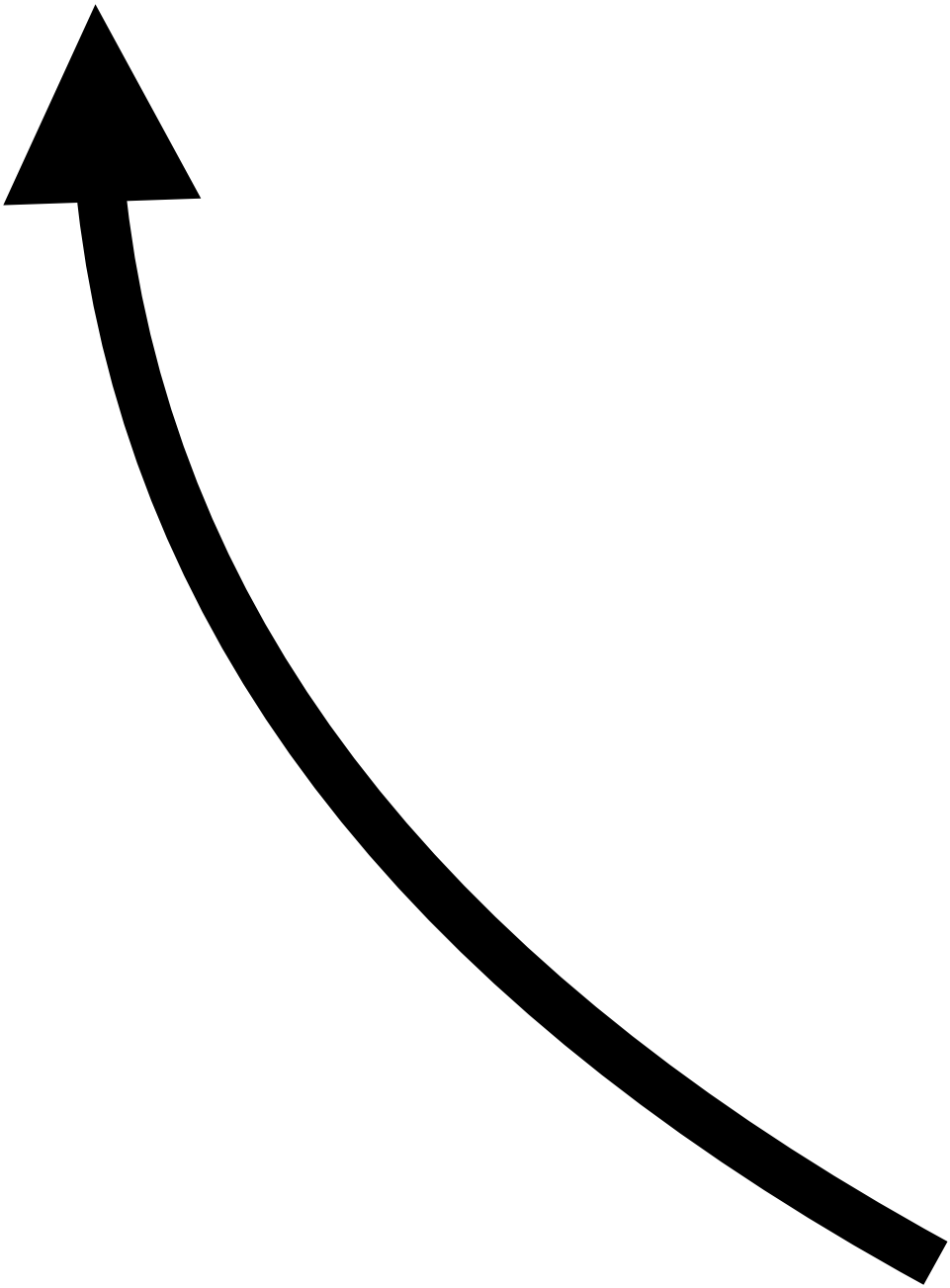
State, Reward



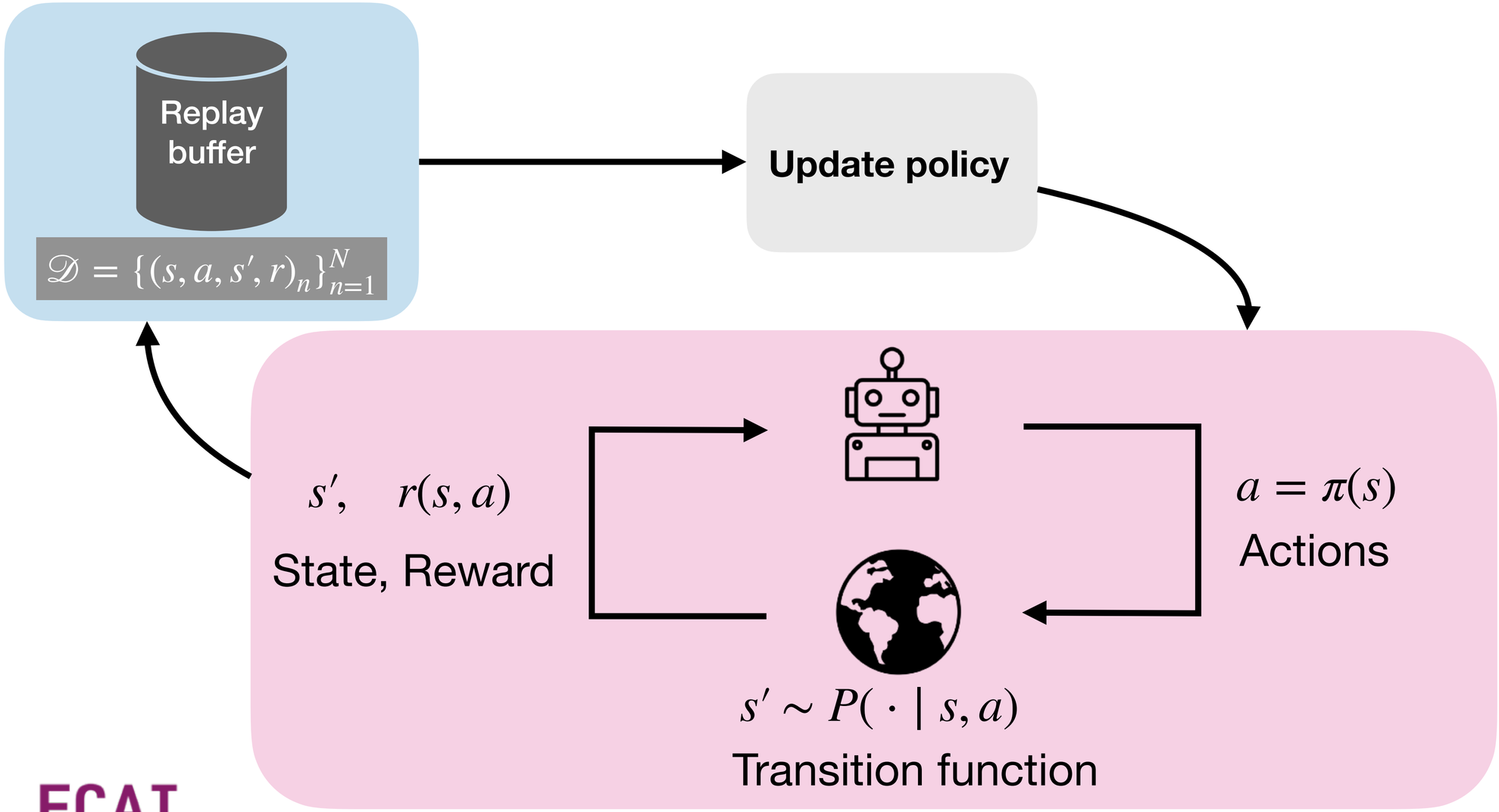
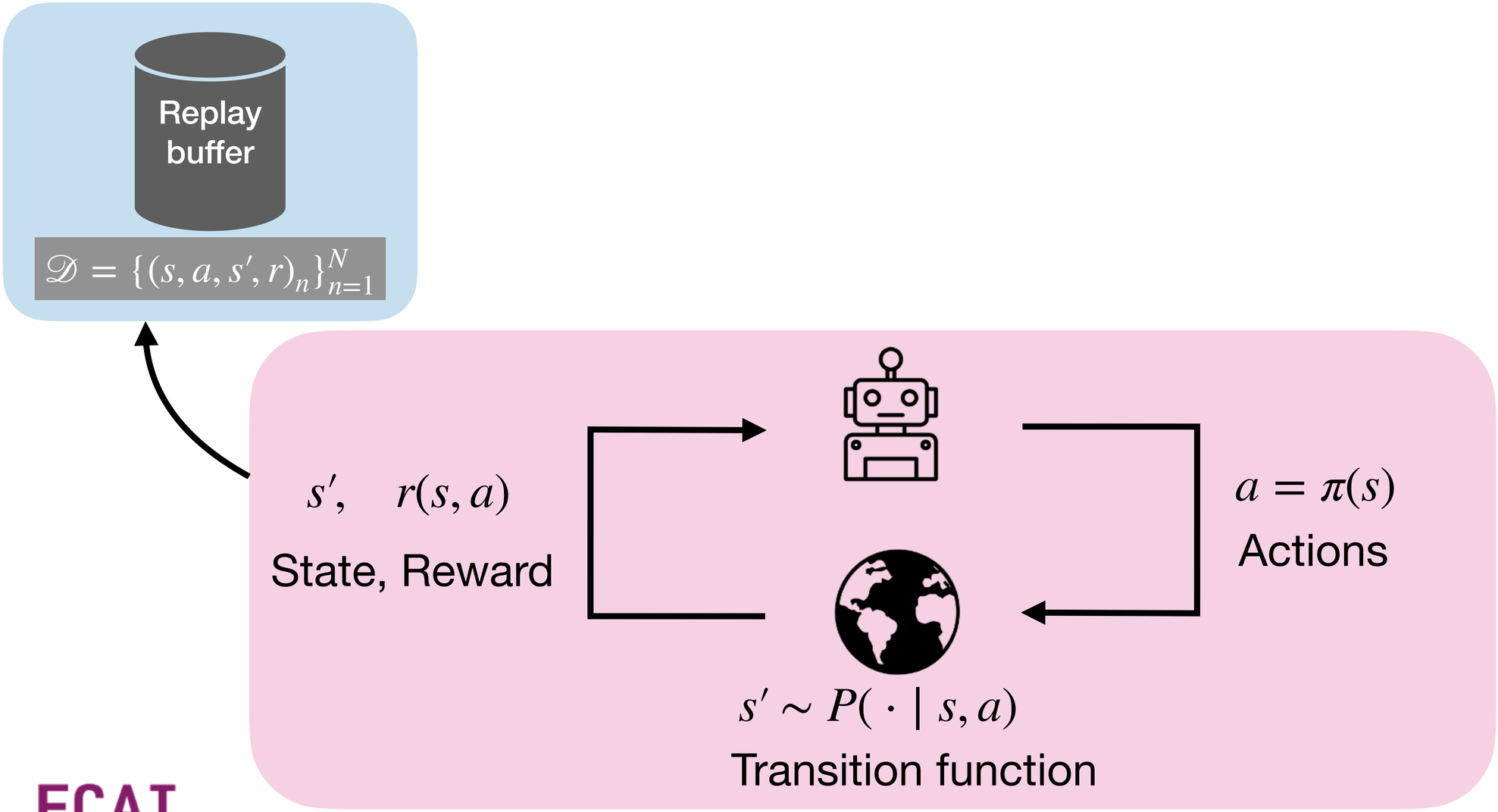
$$\mathscr{D} = \{(s, a, s', r)_n\}_{n=1}^{N}$$

Update policy





Reinforcement Learning



Model-based Reinforcement Learning