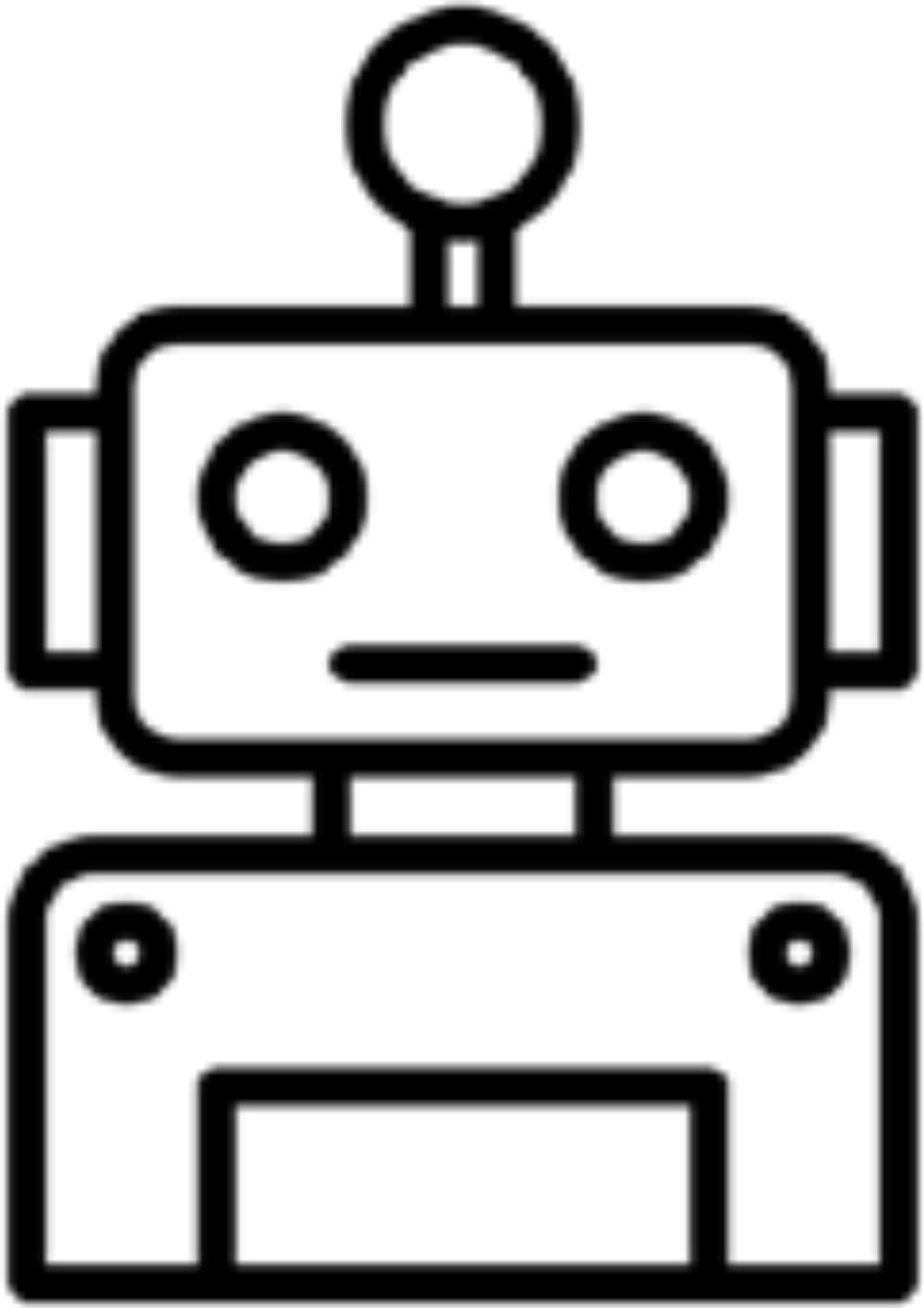
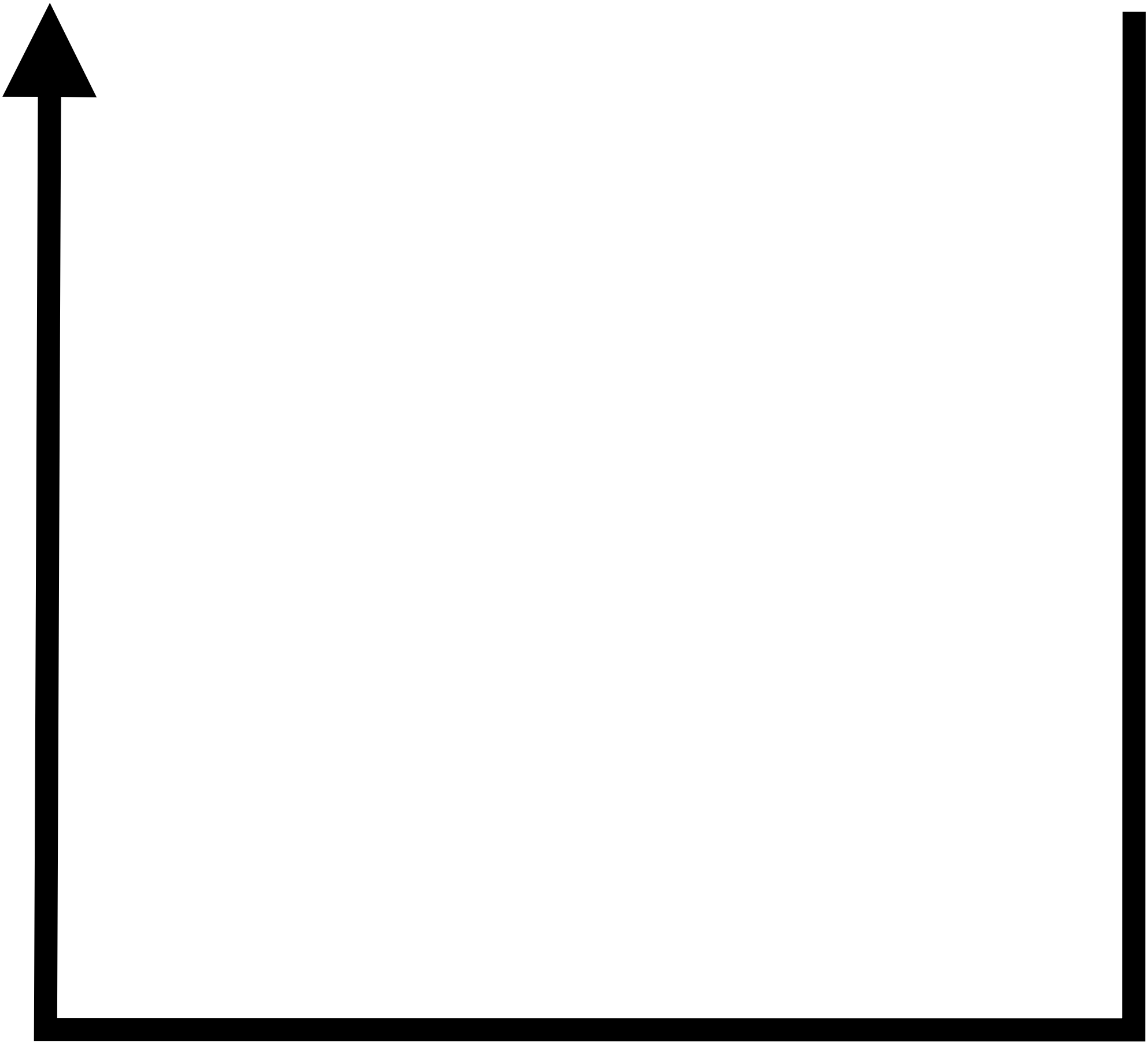
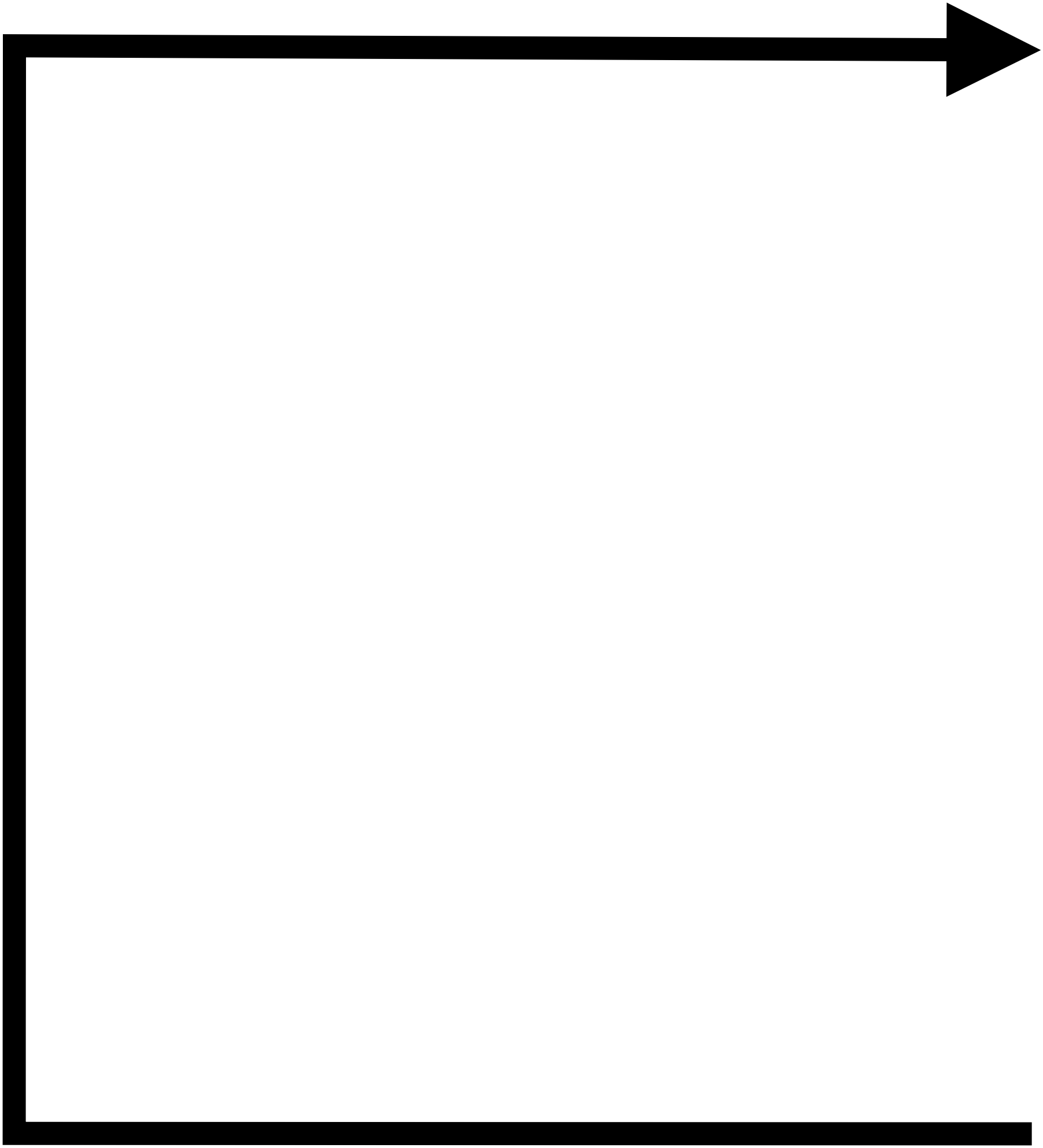
Reinforcement Learning

5







$$a_t = \pi(s_t)$$

Actions

$$s_{t+1} \sim P(\cdot \mid s_t, a_t)$$

Transition function

s_{t+1}, $r(s_t, a_t)$

State, Reward

States $s \in \mathcal{S}$

Actions $a \in \mathcal{A}$

Transition function $P(s_{t+1} \mid s_t, a_t)$

Reward function $r_t = r(s_t, a_t)$

Start state s_0

Discount factor $\gamma \in [0,1]$

Policy $\pi : \mathcal{S} \rightarrow \mathcal{A}$

Markov Decision Process (MDP)

$$s_{t+1} = f(s_t, a_t) + \epsilon_t$$

Reinforcement Learning

Markov Decision Process (MDP)

States $s \in \mathcal{S}$

Actions $a \in \mathcal{A}$

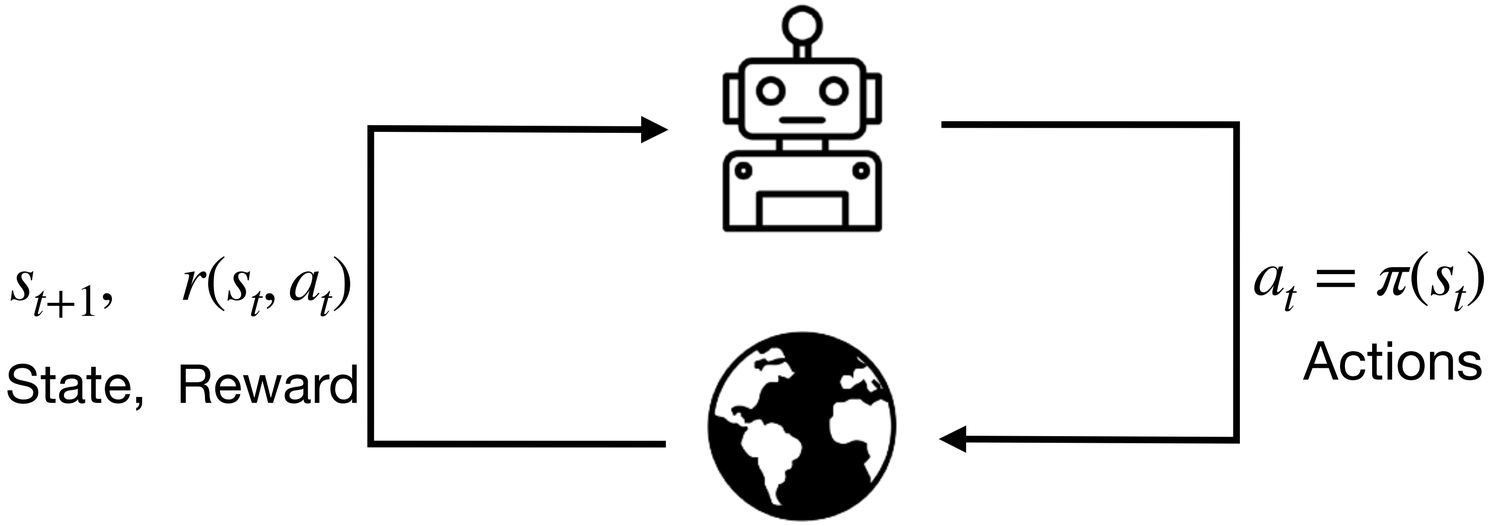
Transition function $P(s_{t+1} \mid s_t, a_t)$

Reward function $r_t = r(s_t, a_t)$

Start state s_0

Discount factor $\gamma \in [0,1]$

Policy $\pi : \mathcal{S} \to \mathcal{A}$



$$s_{t+1} \sim P(\cdot \mid s_t, a_t)$$

Transition function

$$s_{t+1} = f(s_t, a_t) + \epsilon_t$$

Reinforcement Learning

Markov Decision Process (MDP)

States $s \in \mathcal{S}$

Actions $a \in \mathcal{A}$

Transition function $P(s_{t+1} \mid s_t, a_t)$

Reward function $r_t = r(s_t, a_t)$

Start state s_0

Discount factor $\gamma \in [0,1]$

Policy $\pi : \mathcal{S} \rightarrow \mathcal{A}$