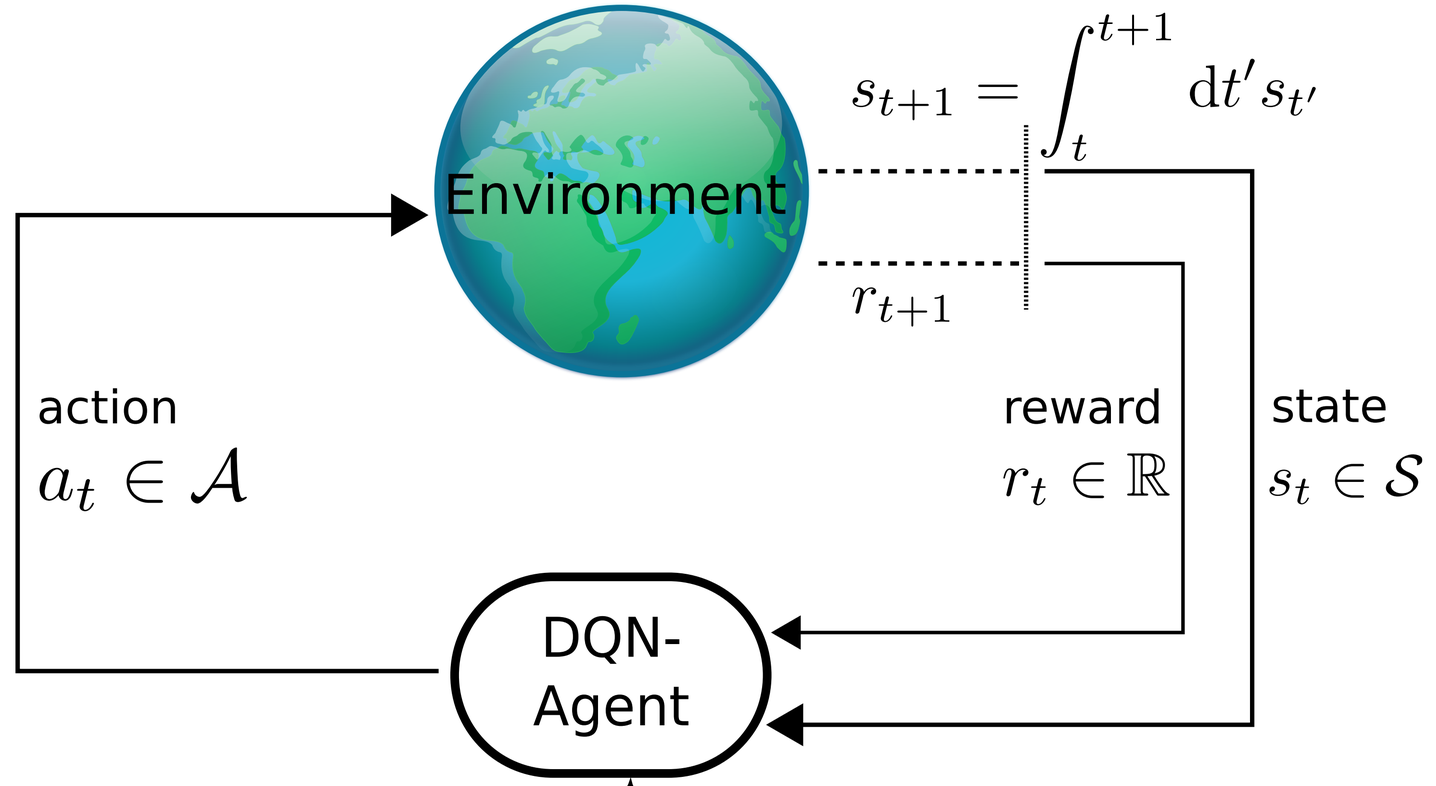
Management Options:
- Carbon Tax
- Subsidies for renewables
- Nature protection policy

⟶ Defines action set $\mathcal{A}$



Reward functions

- Survival:

$$r_t = \begin{cases} 1 & s_t \in \mathcal{O} \\ 0 & \text{else} \end{cases},$$

- Boundary distance:

$$\mathcal{N}_t = \mathbb{L}^2(s_t - SB),$$

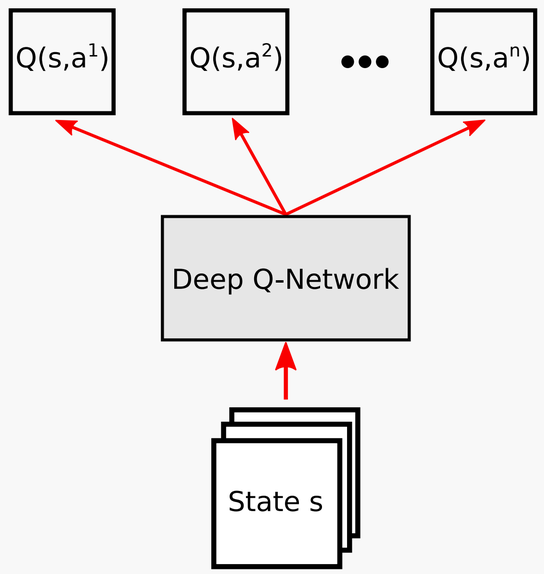
$$r_t = \begin{cases} \mathcal{N}_t & s_t \in \mathcal{O} \\ 0 & \text{else} \end{cases},$$

Deep Reinforcement Learning

Q Learning

$$\underbrace{Q(s,a)}_{\text{Q target}} = \underbrace{r_t(s,a)}_{\substack{\text{Reward of taking} \\ \text{action a} \\ \text{at state s}}} + \gamma \underbrace{\max_{a' \in \mathcal{A}} Q(s_{t+1}, a')}_{\substack{\text{Discounted max} \\ \text{Q-value among all} \\ \text{actions from next} \\ \text{state } s_{t+1}}}$$

Deep Q-Networks



Experience Replay