

Management Options:

- Carbon Tax
- Subsidies for renewables
- Nature protection policy

→ Defines action set \mathcal{A}

action

$$a_t \in \mathcal{A}$$

Environment

$$s_{t+1} = \int_t^{t+1} dt' s_{t'}$$

$$r_{t+1}$$

reward

$$r_t \in \mathbb{R}$$

state

$$s_t \in \mathcal{S}$$

Reward functions

- Survival:

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

- Boundary distance:

$$\mathcal{N}_t = \mathbb{L}^2(s_t - SB),$$
$$r_t = \begin{cases} \mathcal{N}_t & s_t \in \mathcal{O}, \\ 0 & \text{else.} \end{cases}$$

DQN-
Agent

Deep Reinforcement Learning

Q Learning

$$Q(s, a) = r_t(s, a) + \gamma \max_{a' \in \mathcal{A}} Q(s_{t+1}, a')$$

Q target

Reward
of taking
action a
at state s

Discounted max
Q-value among all
actions from next
state s_{t+1}

Deep Q-Networks

$$Q(s, a^1)$$

$$Q(s, a^2)$$

...

$$Q(s, a^n)$$

Deep Q-Network

State s

Experience Replay



Transition
(s_t, a_t, r_t, s_{t+1})

Replay Buffer

1	($s_{\bullet}, a_{\bullet}, r_{\bullet}, s'_{\bullet}$)
2	($s_{\bullet}, a_{\bullet}, r_{\bullet}, s'_{\bullet}$)
	...
k	($s_{\bullet}, a_{\bullet}, r_{\bullet}, s'_{\bullet}$)

Mini-Batch MB

Agent