

Policy

$$\pi : s_t \rightarrow a_t$$

Actions

a_t

Environment

$$f : s_t, a_t \rightarrow r_t, s_{t+1}$$

State

s_{t+1}, r_t

