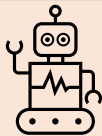


Adapt parameters

η



Reinforcement Learning Agent

rewards

Environment



Actions

Optimization formulation

$$\Delta U_t^* = \underset{\Delta U_t}{\operatorname{argmin}} \sum_t^{t+H-1} \frac{E_t(\Delta U_t)}{E_{\text{avg}}} - \eta \times \frac{\text{Comfort}_{\text{avg}}(T_t(\Delta U_t))}{\text{Comfort}_{\text{avg}}}$$

$$\text{Comfort}(T_t(\Delta U_t)) \geq \mu$$

ΔU_t^*