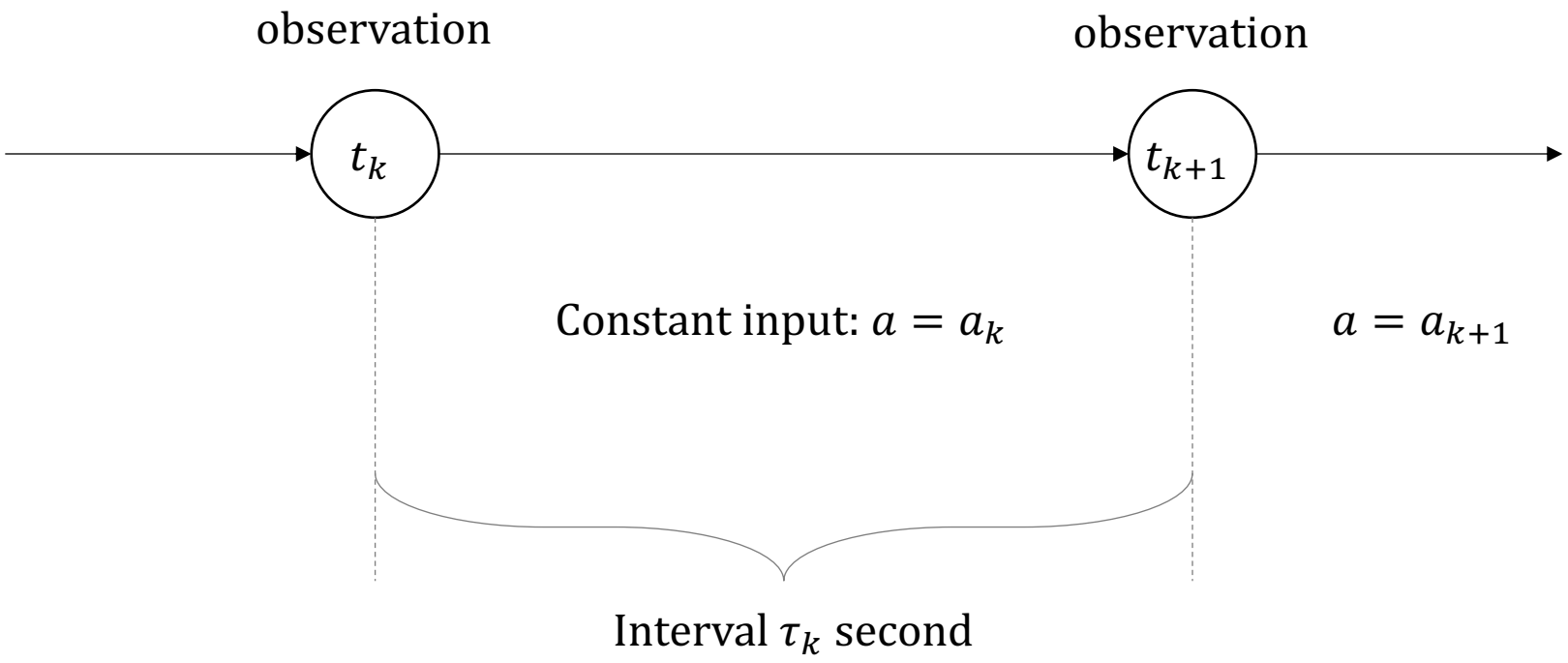
Weekly Report

M2 Ibuki Takeuchi

- Self-triggered control (cf. : Event-triggered control)
  - Controller decides action signal and next observation time
  - (State feedback) control low function: $\pi(s) = [a(s) \quad \tau(s)]$



Weekly Report

M2 Ibuki Takeuchi

- Reinforcement learning for optimal self-triggered control π^*

$$\pi^*(s) = \mathrm{argmax}_\pi J(\pi)$$

$$J(\pi) = \mathbb{E}_{s_0}[V^\pi(s_0)]$$

$$V^\pi(s) = \sum_{i=0}^{\infty} \gamma^i r_i^\pi$$

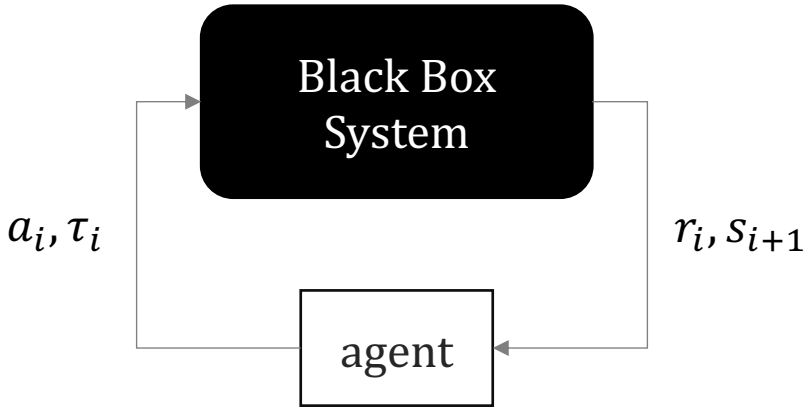
$$r_i^\pi = -\int_{T_i}^{T_{i+1}} s(t)^T Q s(t) dt - \tau_i a_i^T R a_i + \lambda \tau_i$$

- Assume every *i-th* step's access to data tuple (s_t, r_i^π)

Weekly Report

M2 Ibuki Takeuchi

- We can regard this task as a simple RL framework



  - It will converge to optimal policy by long long iteration

- I want to discuss unique feature of this problem for master thesis
  - Ex.) step size ratio of gradient method for $a_\mu(s)$, $\tau_\varphi(s)$

$$\nabla_\mu J(\pi), \nabla_\varphi J(\pi)$$