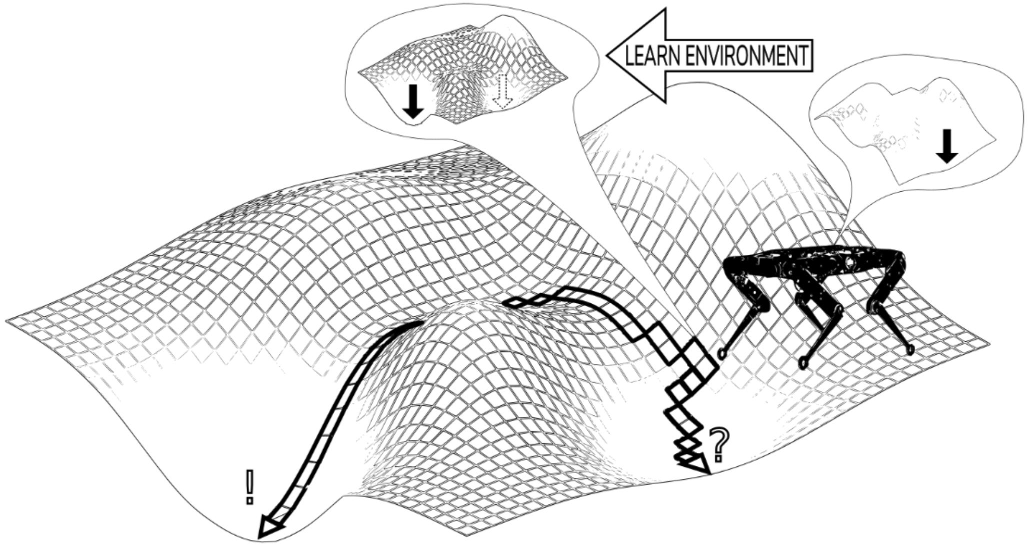
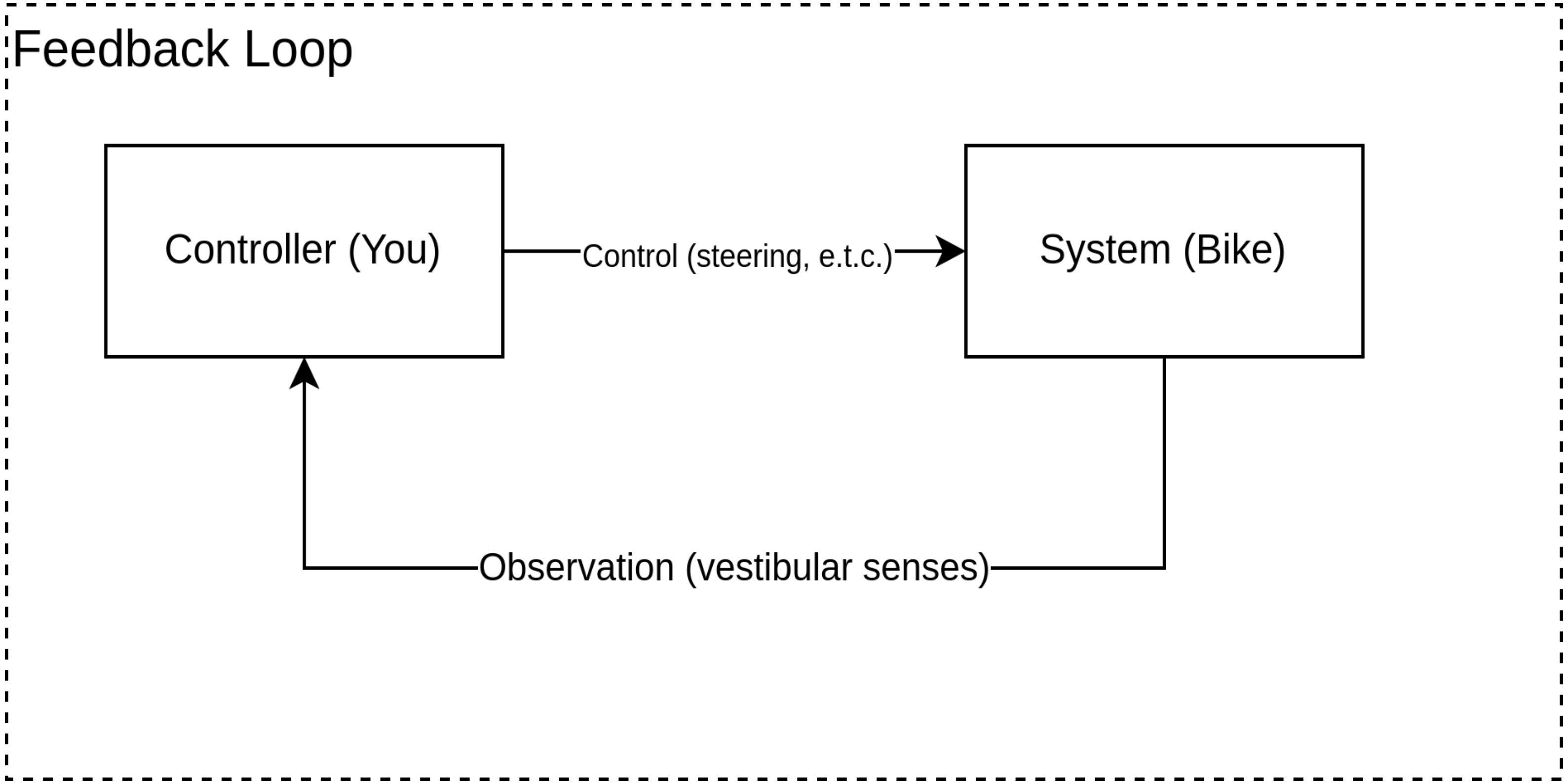
AI in dynamic action



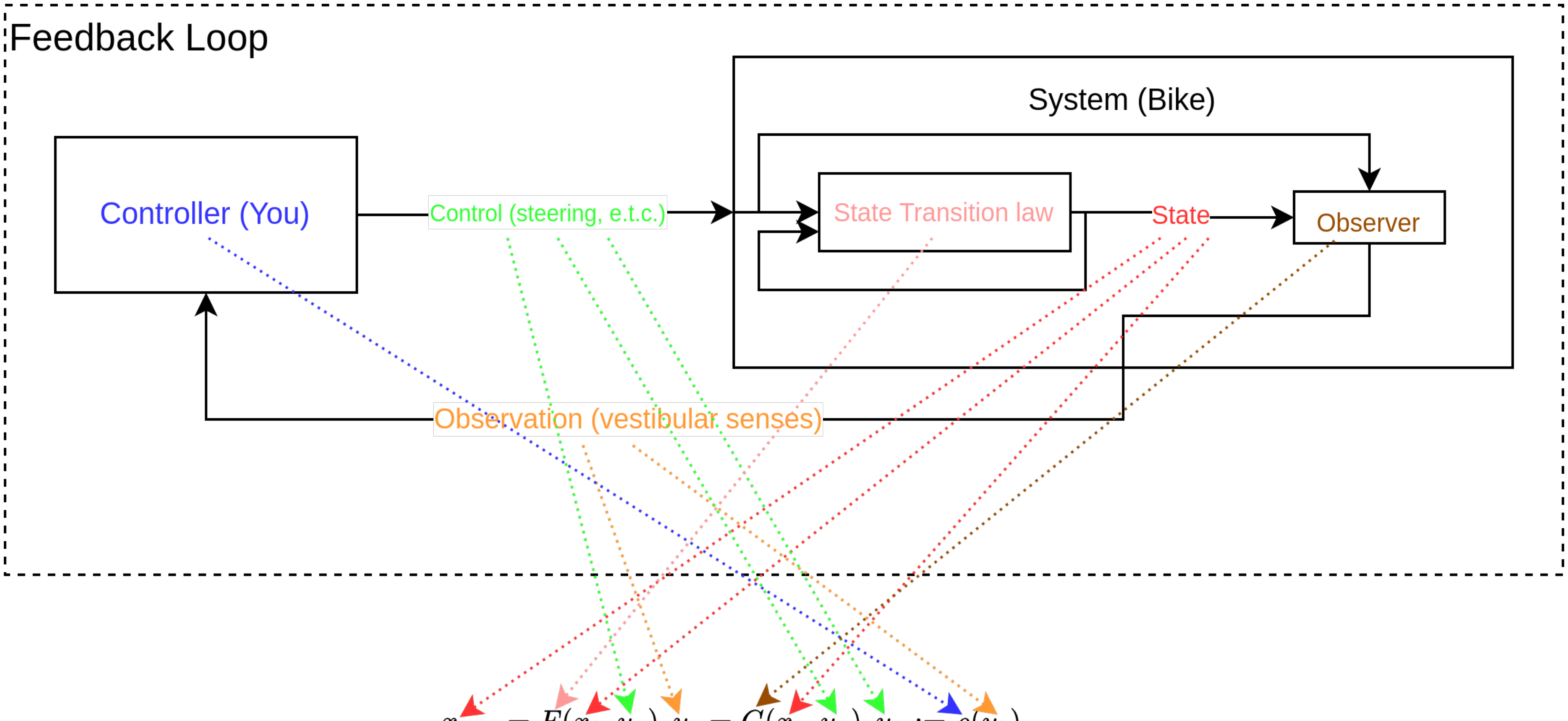
Seminar 1: Terminology

Skolkovo Institute of Science and Technology

Feedback control



Feedback control



$$\underbrace{x_{k+1} = F(x_k, u_k), y_k = G(x_k, u_k)}_{\text{System}}, \underbrace{u_k := \rho(y_k)}_{\text{Control policy}}$$

Please, split into groups of three.

Problems related terminology

System

Objective

Minimize the time it takes to reach your destination.

Discrete time vs. Continuous time

System

Discrete time

$$T := \mathbb{Z}$$

Continuous time

$$T := \mathbb{R}$$

Discrete time vs. Continuous time

$$f : \mathbb{R}^n \times \mathbb{U} \to \mathbb{R}^n$$

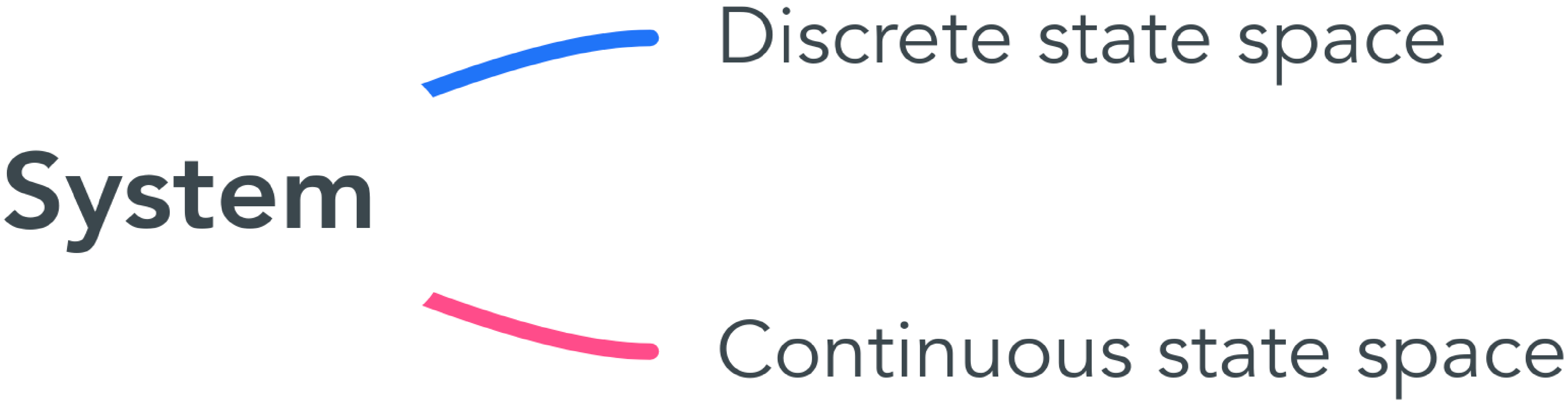
State dynamics function

Continuous: $$\frac{\partial}{\partial t} x(t) = f(x(t), u(t))$$

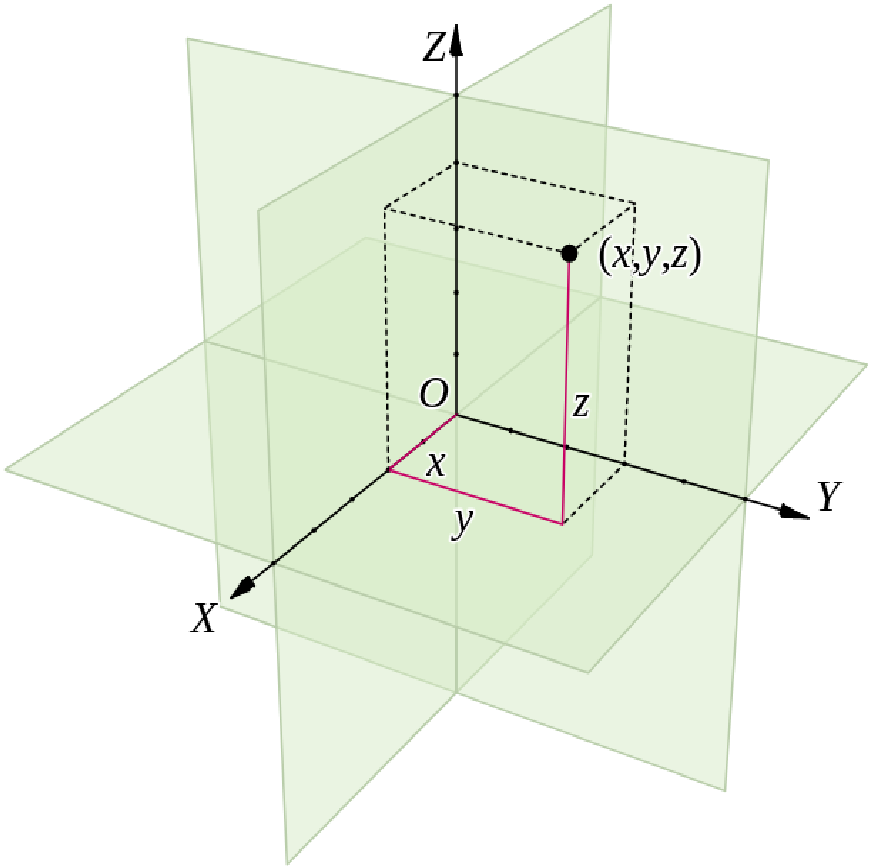
Discrete: $$x_{t+1} = f(x_t, u_t)$$

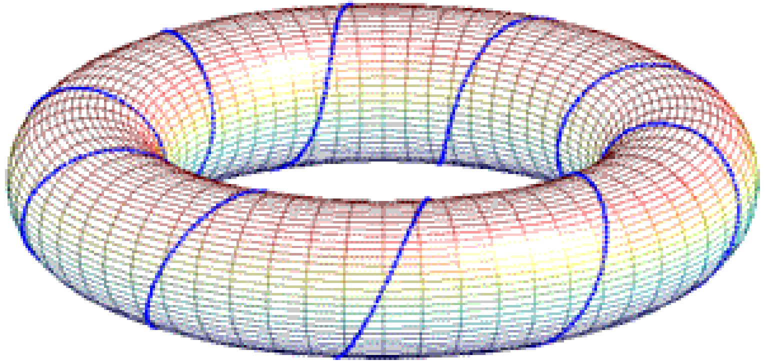
State transition function

Discrete state space vs. Continuous state space



Continuous state space

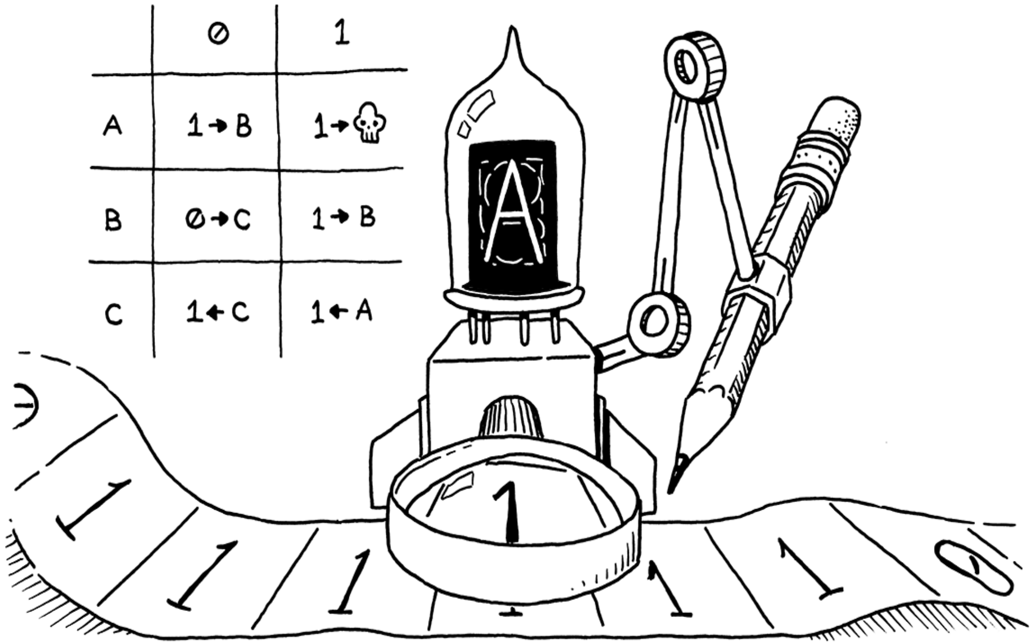




Discrete state space

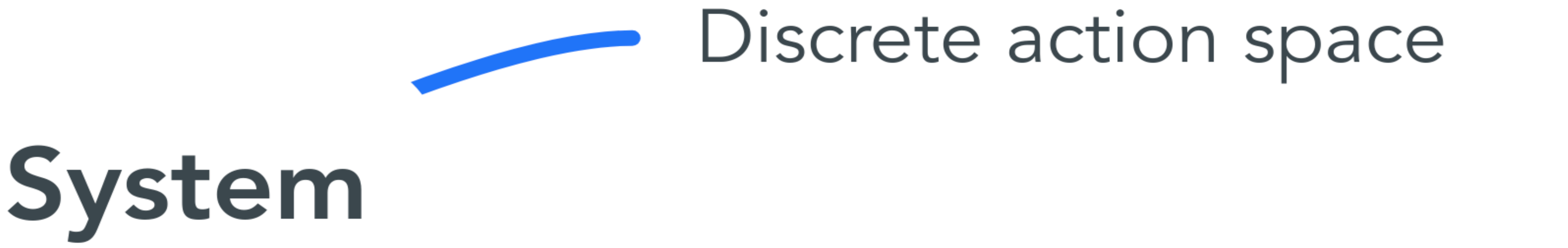
?

Discrete state space



$$x_{t+1} = f(x_t, u_t),$$
$$f : \mathbb{Z} \times \mathbb{U} \to \mathbb{Z}$$

Discrete action space vs. Continuous action space

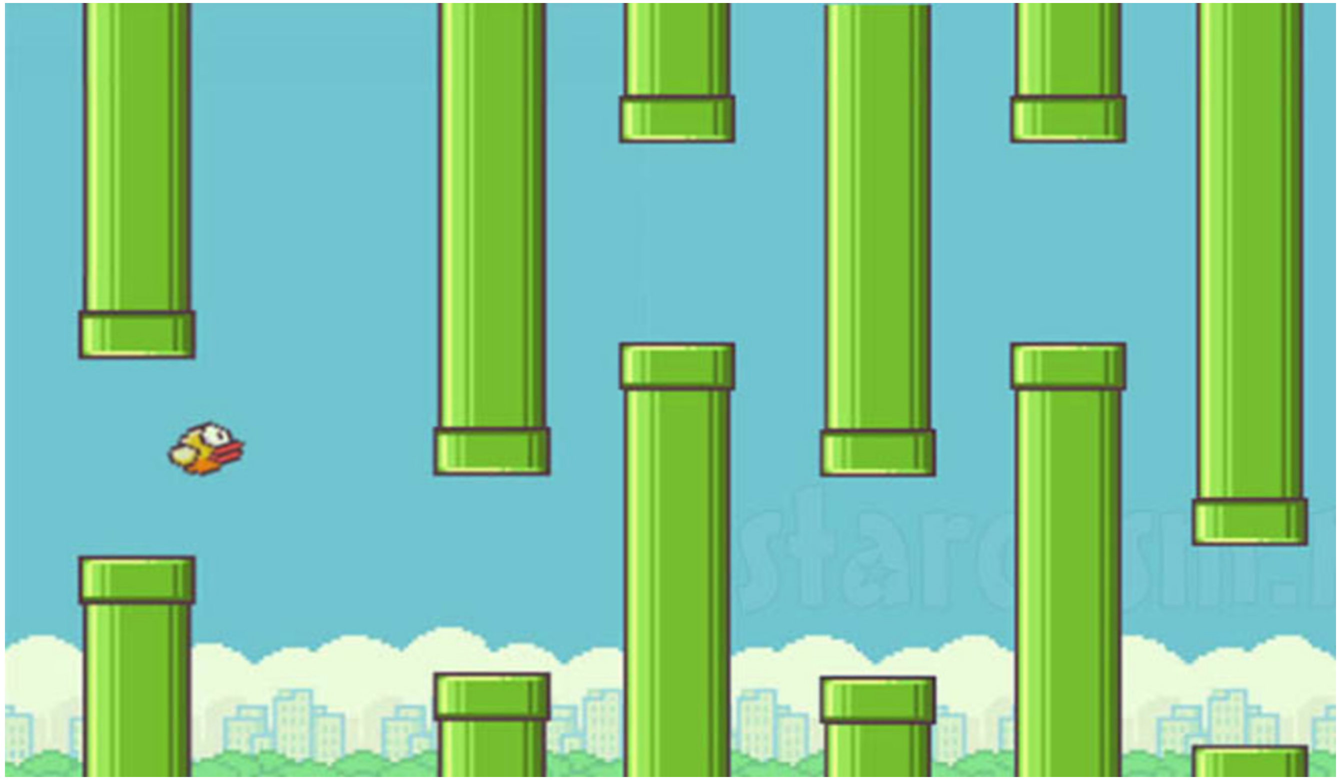


Discrete action space vs. Continuous action space

Continuous

Discrete

Discrete action space with continuous time and state space



Continuous action space with discrete time and state space

?

Continuous action space with discrete time and state space

$$x_{t+1} = f(x_t, u_t) + \lceil \sigma(x_t, u_t)\xi_t \rceil,$$
$$f : \mathbb{Z} \times \mathbb{U} \to \mathbb{Z},$$
$$\sigma : \mathbb{Z} \times \mathbb{U} \to \mathbb{R},$$
$$\xi_t \sim \mathcal{N}(0, 1)$$

Stochastic vs. Deterministic

Stochastic systems

?

Stochastic systems

Pretty much anything

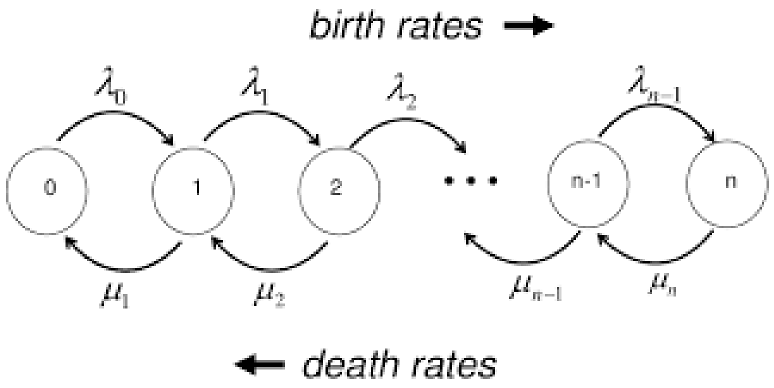
Stochastic systems

Continuous state space

Discrete state space

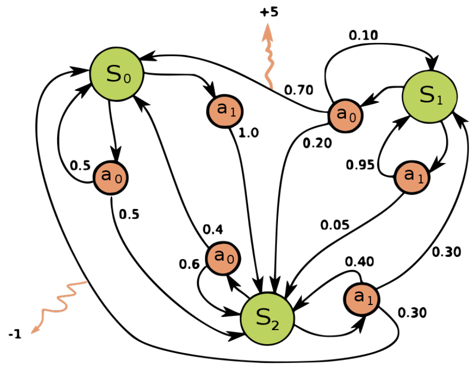
Continuous time

$$\mathrm{d}X_t = f(X_t, U_t)\,\mathrm{d}t + \sigma(X_t, U_t)\,\mathrm{d}W_t,$$
$$W_t \; - \; \text{semimartingale.}$$



Discrete time

$$x_{t+1} \sim \mathcal{F}(x_t, u_t)$$



Full information vs. Partial information

Full information vs. Partial information

$$\textbf{Observation} \stackrel{?}{=} \textbf{State}$$

Full information vs. Partial information

Full information

$$\textbf{Observation} = \textbf{State}$$

Partial information

$$\textbf{Observation} \neq \textbf{State}$$

Partial information examples

$\rho(\cdot)$ – feedback policy.

Full information

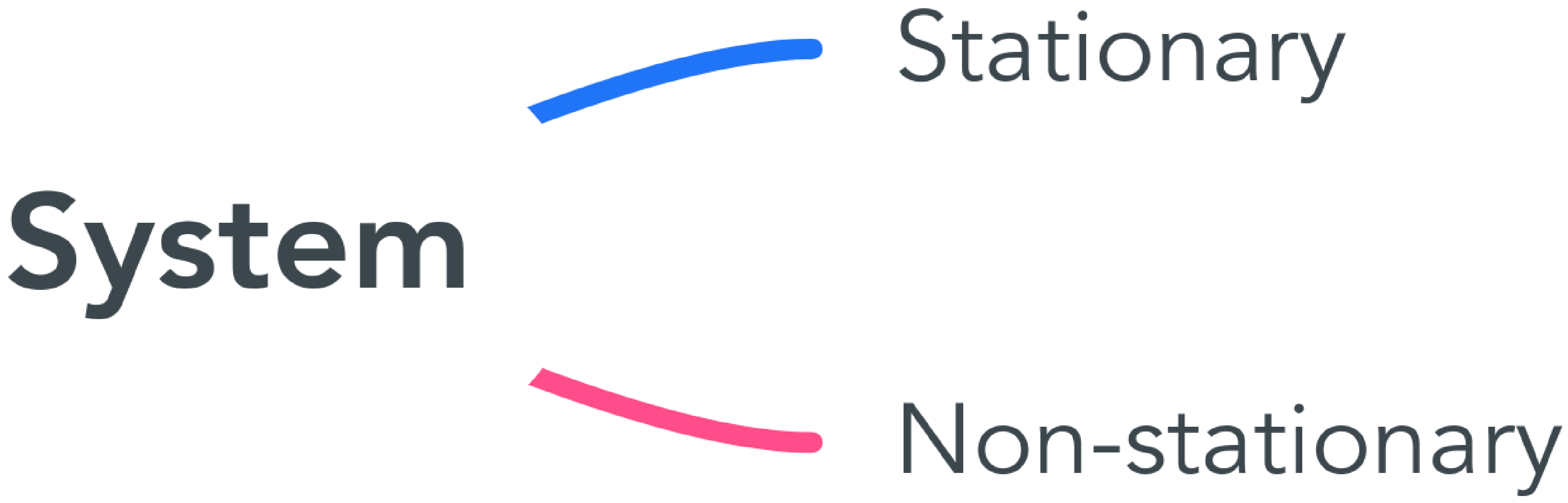
$$u(t) := \rho(x(t))$$

Partial information

$$u(t) := \rho(g(x(t)))$$

$$u(t) := \rho(x(t) + \xi_t), \ \xi \sim \mathcal{N}(\mu, \sigma^2)$$

Stationary vs. Non-stationary



Stationary vs. Non-stationary

Non-stationary

$$x_{t+1} := f(x_t, u_t, t)$$

Stationary

$$x_{t+1} = f(x_t, u_t)$$

Non-stationary --> stationary

$$x_{t+1} := f(x_t, u_t, t) \longrightarrow \begin{aligned} x_{t+1} &= f(x_t, u_t, y_t) \\ y_{t+1} &= y_t + 1 \end{aligned}$$

Example of a non-stationary system

?

Example of a non-stationary system

Cost vs. Reward

Cost vs. Reward

Cost $\longrightarrow$ Minimize

Reward $\longrightarrow$ Maximize

Finite horizon vs. Infinite horizon

Finite horizon vs. Infinite horizon

Finite-horizon

You optimize the objective over a finite time frame.

Infinite-horizon

You optimize the objective over an infinite time frame. (As if your RL agent were to run for all eternity)

Running vs. Terminal

Running vs. Terminal

$J(\cdot,\cdot)$ – total objective.

Running objective

$$J(x(\cdot), u(\cdot)) := \int_{t_1}^{t_2} r(x(t), u(t))\,\mathrm{d}t + T(x(t_2))$$

$$J(x_\cdot, u_\cdot) := \sum_{i=t_1}^{t_2} r(x_i, u_i) + T(x_{t_2})$$

Terminal objective

Terminal objective

Running objective

?

Running objective

Running --> Terminal

$$\mathrm{x}_{t+1} = f(x_t, u_t),$$
$$J = \sum_{i=0}^{T} r(x_i, u_i)$$

$$\longrightarrow$$

$$\mathrm{x}_{t+1} = f(x_t, u_t),$$
$$j_{t+1} = j_t + r(x_t, u_t),$$
$$J = j_T$$

Terminal --> Running

$$\mathrm{x}_{t+1} = f(x_t, u_t),$$
$$J = T(x_T)$$

$$\mathrm{x}_{t+1} = f(x_t, u_t),$$
$$j_{t+1} = T(f(x_t, u_t)) - T(x_t),$$
$$J = \sum_{i=0}^{T} j_i$$

Q&A