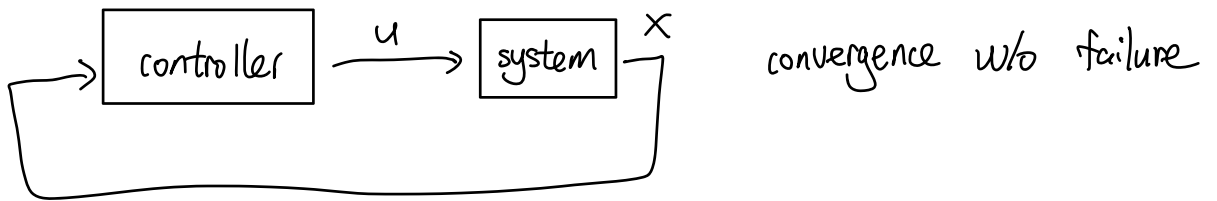


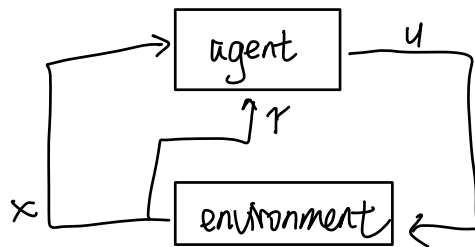
<u>Control</u>		<u>RL</u>
system		environment
controller		agent
dynamics		transition
state (x)		s
control input	u	a
policy	$\pi(\cdot)$	$\pi(\cdot)$

Adaptive control

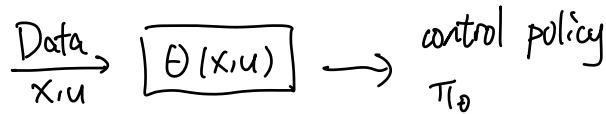
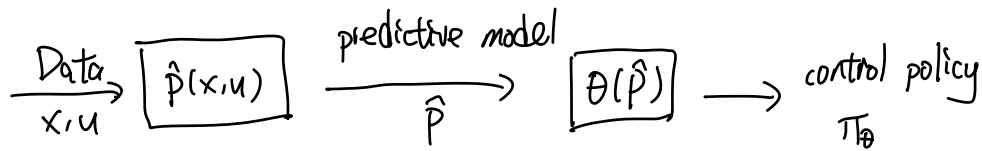


$$\max_{\theta} \mathbb{E}_d J(x_{x,t}^{\tau_{\theta,d}}) \quad \text{performance objective}$$

RL



Learn from failure
 \Rightarrow few assumptions



Direct adaptive control

Model-free RL

exploration
vs
exploitation

Safe learning problem?

use a safety filter (aka safety critic) to find a safe region for learning.

- require privileged information during training
- naive safety filter
- assumptions used for safety analysis may not hold