Three main ingredients

- Mathematical description of the system to be controlled (state-space models)
  - ➡ We use MDPs
- Specification of a performance criterion (the cost function)
  - ➡ The reward designed by us (or emitted by the environment in RL setting)
- Specification of constraints
  - ➡ Control or state constraints

Model assumptions

- Discrete time
- A stochastic dynamics with Markov property:
  $\mathbf{s}_{t+1} = \mathbf{f}(\mathbf{s}_t, \mathbf{a}_t, \omega_t)$ with $\omega_t = \omega_{t-1}(\mathbf{s}_t, \mathbf{a}_t)$
- Later ω_t is normally distributed
- In stochastic settings optimise for an expected reward