## Reinforcement learning

| | | |
|---|---|---|
| $s, s'$ | State | $s_t, s_{t+1}$ |
| $a$ | Action | $a_t$ |
| $p$ | State transition probability | $s' \sim p\left(s' \vert s, a\right)$ |
| $r$ | Reward | $r_t = r(s_t, a_t)$ |
| $\pi$ | Policy | $a \sim \pi(a \vert s)$, $a = \pi(s)$ |
| $\gamma$ | Discount factor | $\gamma \in [0, 1]$ |
| $G_t$ | Discounted return | $G_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k}$ |
| $V^\pi$ | Value function | $V^\pi(s) = \mathbb{E}\left[G_0 \vert s_0 = s\right]$ |
| $V^\star$ | Optimal value function | The above, but better |
| $Q^\pi$ | State-action value function | |
| $Q^\star$ | Optimal value function | |

## Control

| | | |
|---|---|---|
| $x$ | State | $x_t$ |
| $u$ | (Control) input | $u_t$ |
| $f$ | State transition function | $x_{t+1} = f(x_t, u_t)$ |
| $\ell$ | (Stage) cost | $\ell(x, u) = x^T M x + u^T R u$ |
| $K$ | Gain matrix | $u = -Kx$ |

## Acronyms

| | |
|---|---|
| RL | Reinforcement learning |
| MPC | Model predictive control |
| LQR | Linear quadratic regulator |
| PID | Proportional-integral-derivative |