Appendix A: Notation and correspondence between RL and control

Throughout this tutorial we adopt a unified notation whenever possible. Since RL and control theory have historically developed different terminologies and conventions, Table 1 summarizes the main symbols used throughout the text and their correspondence with common notation in each community.

Table 1 Correspondence between the notation adopted in this tutorial and common notation in RL and control theory.

Concept	Tutorial	Common RL	Common Control
State	s_t	s_t	x_t
Action / control input	a_t	a_t	u_t
Policy / controller	π	π	κ, $u = \kappa(x)$
Parameterized policy	π_θ	π_θ	κ_θ
Dynamics	$f(s,a)$ or $p(\cdot\|s,a)$	P, p	f, A, B
Disturbance	η_t	often implicit	d_t, w_t, η_t
State space	$\mathcal{S}$	$\mathcal{S}$	$\mathcal{X}$
Action space	$\mathcal{A}$	$\mathcal{A}$	$\mathcal{U}$
Cost	$c(s,a)$	$-r(s,a)$	$\ell(x,u)$
Reward	$r(s,a)$	$r(s,a)$	rarely used
Objective (min.)	$J(\pi)$	uncommon	$J(\pi)$
Objective (max.)	$V(\pi)$	$J(\pi)$	uncommon
Discount factor	γ	γ	often omitted
Value function	V^π	V^π	cost-to-go
Optimal value function	$V^\star$	$V^\star$	optimal cost-to-go
Action-value function	Q^π	Q^π	state-action cost-to-go
Lyapunov function	rarely used	U	V
Advantage function	A^π	A^π	rarely used
State distribution	d^π	d^π	invariant measure
Actor	π_θ	π_θ	parameterized controller
Critic	q_w	Q_w, Q_ϕ	approximate cost-to-go
Reference signal	r_t	goal signal	r_t
Tracking error	e_t	uncommon	e_t
Feedback gain	K	uncommon	K