

Task	Methods					
	TRA	GCBC	CRL	GCIQL	GCIVL	QRL
antmaze medium stitch	60.7 (± 3.0)*	45.5(± 3.9)	52.7(± 2.2)	29.3(± 2.2)	44.1(± 2.0)	59.1(± 2.4)
antmaze large stitch	12.8 (± 2.0)	3.4(± 1.0)	10.8(± 0.6)	7.5(± 0.7)	18.5 (± 0.8) [†]	18.4 (± 0.7)
antsoccer arena stitch	17.0(± 1.2)	24.5 (± 2.8)	0.7(± 0.1)	2.1(± 0.1)	21.4(± 1.1)	0.8(± 0.2)
humanoidmaze medium stitch	46.1 (± 1.9)	29.0(± 1.7)	36.2(± 0.9)	12.1(± 1.1)	12.3(± 0.6)	18.0(± 0.7)
humanoidmaze large stitch	8.6 (± 1.4)	5.6(± 1.0)	4.0(± 0.2)	0.5(± 0.1)	1.2(± 0.2)	3.5(± 0.5)
antmaze large navigate	35.4 (± 1.8)	24.0(± 0.6)	82.8 (± 1.4)	34.2(± 1.3)	15.7(± 1.9)	74.6(± 2.3)
cube single noisy	9.2(± 0.9)	8.4(± 1.0)	38.3(± 0.6)	99.3 (± 0.2)	70.6(± 3.3)	25.5(± 2.1)

RL methods with a separate value network to update the actor are in gray.

*The best non-RL methods up to significance are **highlighted**.

[†]We **bold** the best performance across all methods.

