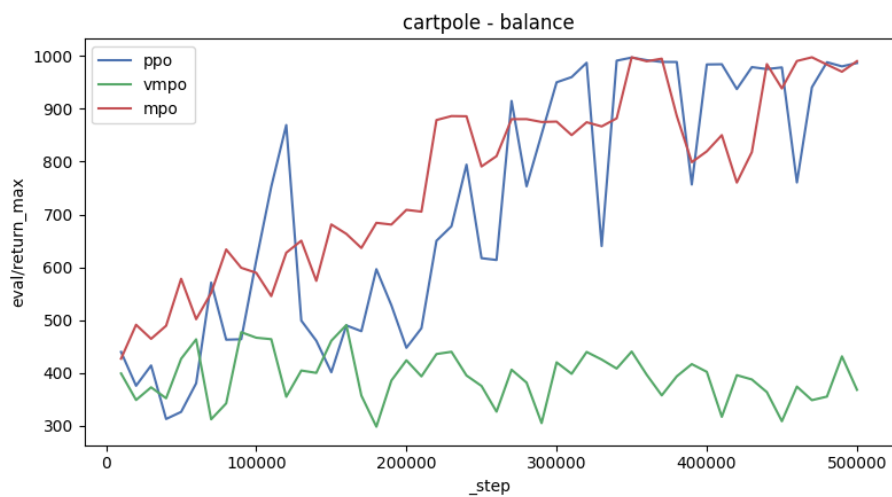


## Report

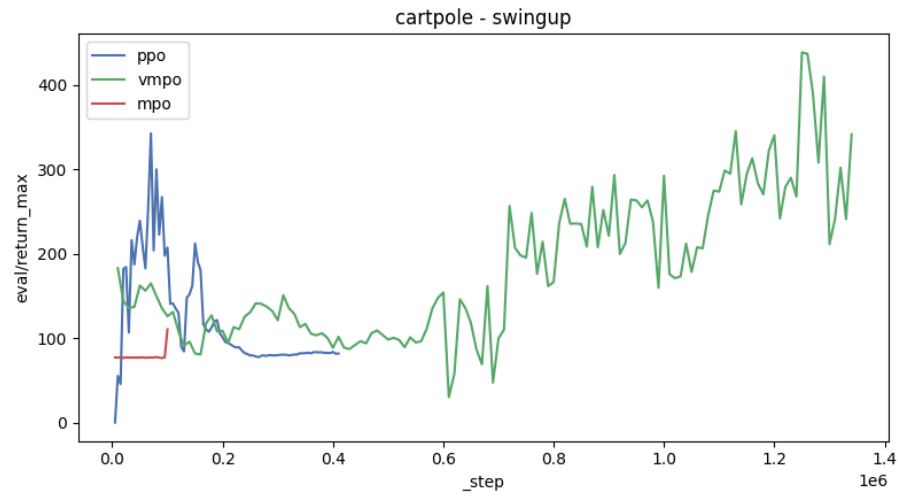
Domain	Task	ppo	vmppo	mpo
cartpole	balance	986	368	990
cartpole	swingup	82	341	110
cheetah	run	368	282	301
walker	run	44	100	488
walker	walk	456	245	857

### cartpole - balance



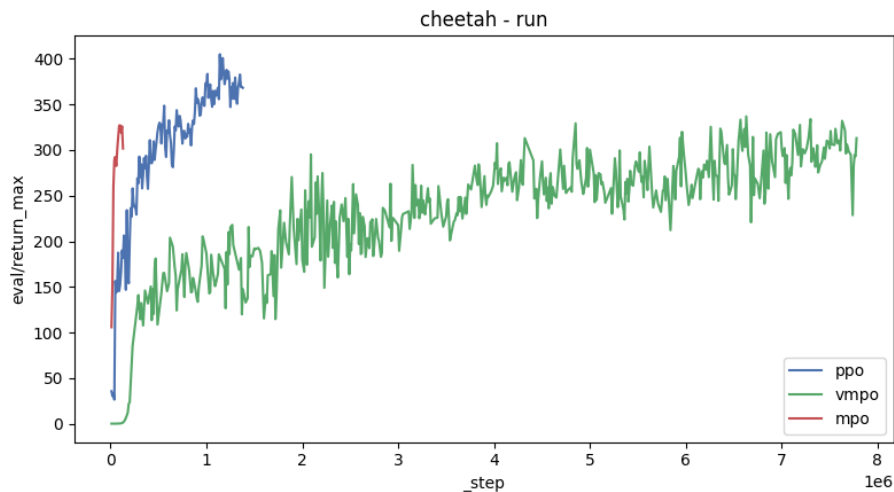
Run	Algorithm	_step	eval/return_max
mpo-cartpole-balance	mpo	500000	990
mpo-cartpole-balance	mpo	111461	436
ppo-cartpole-balance	ppo	500000	986
vmppo-cartpole-balance	vmppo	500000	368
vmppo-cartpole-balance	vmppo	500000	309
vmppo-cartpole-balance	vmppo	500000	155
vmppo-cartpole-balance	vmppo	500000	-

## cartpole - swingup



Run	Algorithm	_step	eval/return_max
mpo-cartpole-swingup	mpo	100638	110
mpo-cartpole-swingup	mpo	100405	34
mpo-cartpole-swingup	mpo	80000	30
mpo-cartpole-swingup	mpo	106228	29
mpo-cartpole-swingup	mpo	101164	0
mpo-cartpole-swingup	mpo	57000	0
ppo-cartpole-swingup	ppo	415000	82
vmppo-cartpole-swingup	vmppo	1349632	341
vmppo-cartpole-swingup	vmppo	593920	207
vmppo-cartpole-swingup	vmppo	560000	135
vmppo-cartpole-swingup	vmppo	800000	93

## cheetah - run

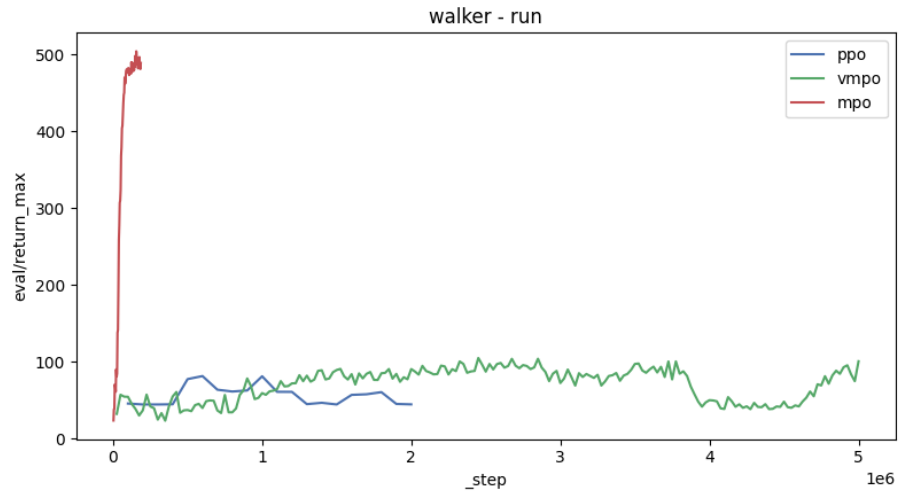


Run	Algorithm	_step	eval/return_max
mpo-cheetah-run	mpo	130855	301
mpo-cheetah-run	mpo	155658	280
mpo-cheetah-run	mpo	500000	70
mpo-cheetah-run	mpo	229137	0
ppo-cheetah-run	ppo	1390000	368
ppo-cheetah-run-GOOD	ppo	3000000	346
ppo-cheetah-run	ppo	1786000	187
ppo-cheetah-run	ppo	1000000	162
ppo-cheetah-run	ppo	3000000	114
ppo-cheetah-run	ppo	479232	106
ppo-cheetah-run	ppo	960000	85
ppo-cheetah-run	ppo	500000	82
ppo-cheetah-run	ppo	1000000	76
ppo-cheetah-run	ppo	500000	16
ppo-cheetah-run	ppo	500000	14
ppo-cheetah-run	ppo	310000	14
vmppo-cheetah-run	vmppo	7800000	282
vmppo-cheetah-run	vmppo	8492000	238
vmppo-cheetah-run	vmppo	1802240	238
vmppo-cheetah-run	vmppo	1980000	223
vmppo-cheetah-run	vmppo	812000	215
vmppo-cheetah-run	vmppo	2481000	200
vmppo-cheetah-run	vmppo	1564672	200
vmppo-cheetah-run-OKAY	vmppo	4681728	196
vmppo-cheetah-run	vmppo	1081000	196

Run	Algorithm	_step	eval/return_max
vmpo-cheetah-run	vmpo	2473000	188
vmpo-cheetah-run	vmpo	1605000	184
vmpo-cheetah-run	vmpo	5278000	178
vmpo-cheetah-run	vmpo	1721000	177
vmpo-cheetah-run	vmpo	3389000	176
vmpo-cheetah-run	vmpo	666665	172
vmpo-cheetah-run	vmpo	1108000	171
vmpo-cheetah-run	vmpo	3500000	168
vmpo-cheetah-run	vmpo	1979976	161
vmpo-cheetah-run	vmpo	828000	155
vmpo-cheetah-run	vmpo	3522560	153
vmpo-cheetah-run	vmpo	569000	150
vmpo-cheetah-run	vmpo	609000	145
vmpo-cheetah-run	vmpo	225000	145
vmpo-cheetah-run	vmpo	10000000	145
vmpo-cheetah-run	vmpo	5111000	143
vmpo-cheetah-run	vmpo	1510000	141
vmpo-cheetah-run	vmpo	1276000	138
vmpo-cheetah-run	vmpo	500000	130
vmpo-cheetah-run	vmpo	835000	127
vmpo-cheetah-run	vmpo	406404	126
vmpo-cheetah-run	vmpo	500000	125
vmpo-cheetah-run	vmpo	387000	123
vmpo-cheetah-run	vmpo	1699000	111
vmpo-cheetah-run	vmpo	1750000	109
vmpo-cheetah-run	vmpo	500000	102
vmpo-cheetah-run	vmpo	942080	99
vmpo-cheetah-run	vmpo	960952	96
vmpo-cheetah-run	vmpo	1420000	96
vmpo-cheetah-run	vmpo	540672	95
vmpo-cheetah-run	vmpo	206000	88
vmpo-cheetah-run	vmpo	565248	83
vmpo-cheetah-run	vmpo	384000	80
vmpo-cheetah-run	vmpo	1153000	77
vmpo-cheetah-run	vmpo	1600000	77
vmpo-cheetah-run	vmpo	1706000	63
vmpo-cheetah-run	vmpo	1250000	57
vmpo-cheetah-run	vmpo	50000	43
vmpo-cheetah-run	vmpo	604000	34
vmpo-cheetah-run	vmpo	98304	11
vmpo-cheetah-run	vmpo	500000	10
vmpo-cheetah-run	vmpo	436000	9
vmpo-cheetah-run	vmpo	992960	5
vmpo-cheetah-run	vmpo	2150016	5

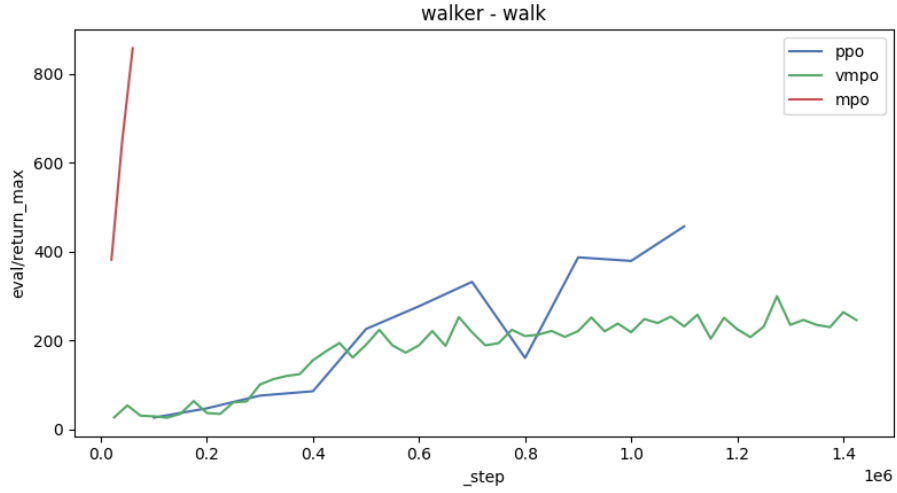
Run	Algorithm	_step	eval/return_max
vmppo-cheetah-run	vmppo	150000	4
vmppo-cheetah-run	vmppo	125000	4
vmppo-cheetah-run	vmppo	1137120	3
vmppo-cheetah-run	vmppo	673000	2
vmppo-cheetah-run	vmppo	1729696	1
vmppo-cheetah-run	vmppo	258256	0
vmppo-cheetah-run	vmppo	95090	0
vmppo-cheetah-run	vmppo	220215	0

### walker - run



Run	Algorithm	_step	eval/return_max
mpo-walker-run	mpo	188111	488
mpo-walker-run	mpo	259819	478
ppo-walker-run	ppo	2091000	44
vmppo-walker-run	vmppo	5000000	100
vmppo-walker-run	vmppo	574000	86
vmppo-walker-run	vmppo	5000000	78
vmppo-walker-run	vmppo	1229000	76
vmppo-walker-run	vmppo	5000000	46
vmppo-walker-run	vmppo	5000000	34

## walker - walk



Run	Algorithm	_step	eval/return_max
mpo-walker-walk	mpo	62152	857
mpo-walker-walk	mpo	131301	739
mpo-walker-walk	mpo	124645	657
mpo-walker-walk	mpo	245772	25
ppo-walker-walk	ppo	1113000	456
ppo-walker-walk	ppo	1265000	439
ppo-walker-walk	ppo	220000	121
ppo-walker-walk	ppo	300000	113
ppo-walker-walk	ppo	1029000	68
ppo-walker-walk	ppo	364544	66
ppo-walker-walk	ppo	243000	46
vmppo-walker-walk	vmppo	1427000	245
vmppo-walker-walk	vmppo	3000000	184
vmppo-walker-walk	vmppo	1329000	158
vmppo-walker-walk	vmppo	3000000	153
vmppo-walker-walk	vmppo	520000	111
vmppo-walker-walk	vmppo	3000000	102
vmppo-walker-walk	vmppo	300000	89
vmppo-walker-walk	vmppo	3890000	64
vmppo-walker-walk	vmppo	350000	39
vmppo-walker-walk	vmppo	511000	34
vmppo-walker-walk	vmppo	74000	29
vmppo-walker-walk	vmppo	52000	29