

2 of 7

```

317782/1500000: episode: 218, duration: 13.599s, episode steps: 2204, steps per second: 162, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
320264/1500000: episode: 219, duration: 15.053s, episode steps: 2482, steps per second: 165, episode reward: -7.000, mean reward: -0.003 [-1.000, 1.000], mean acti
322601/1500000: episode: 220, duration: 14.096s, episode steps: 2337, steps per second: 166, episode reward: -9.000, mean reward: -0.004 [-1.000, 1.000], mean acti
324820/1500000: episode: 221, duration: 13.528s, episode steps: 2219, steps per second: 164, episode reward: -10.000, mean reward: -0.005 [-1.000, 1.000], mean acti
326657/1500000: episode: 222, duration: 11.149s, episode steps: 1837, steps per second: 165, episode reward: -15.000, mean reward: -0.008 [-1.000, 1.000], mean acti
328929/1500000: episode: 223, duration: 13.924s, episode steps: 2272, steps per second: 163, episode reward: -11.000, mean reward: -0.005 [-1.000, 1.000], mean acti
330659/1500000: episode: 224, duration: 10.611s, episode steps: 1730, steps per second: 163, episode reward: -14.000, mean reward: -0.008 [-1.000, 1.000], mean acti
333055/1500000: episode: 225, duration: 14.587s, episode steps: 2396, steps per second: 164, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean acti
335164/1500000: episode: 226, duration: 12.875s, episode steps: 2109, steps per second: 164, episode reward: -11.000, mean reward: -0.005 [-1.000, 1.000], mean acti
337953/1500000: episode: 227, duration: 16.891s, episode steps: 2789, steps per second: 165, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
340192/1500000: episode: 228, duration: 13.604s, episode steps: 2239, steps per second: 165, episode reward: -7.000, mean reward: -0.003 [-1.000, 1.000], mean acti
342776/1500000: episode: 229, duration: 16.033s, episode steps: 2584, steps per second: 161, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
344121/1500000: episode: 230, duration: 8.262s, episode steps: 1345, steps per second: 163, episode reward: -15.000, mean reward: -0.011 [-1.000, 1.000], mean acti
346423/1500000: episode: 231, duration: 13.941s, episode steps: 2302, steps per second: 165, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
348594/1500000: episode: 232, duration: 13.170s, episode steps: 2171, steps per second: 165, episode reward: -11.000, mean reward: -0.005 [-1.000, 1.000], mean acti
350968/1500000: episode: 233, duration: 14.449s, episode steps: 2374, steps per second: 164, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
353014/1500000: episode: 234, duration: 12.646s, episode steps: 2046, steps per second: 162, episode reward: -14.000, mean reward: -0.007 [-1.000, 1.000], mean acti
355191/1500000: episode: 235, duration: 13.347s, episode steps: 2177, steps per second: 163, episode reward: -12.000, mean reward: -0.006 [-1.000, 1.000], mean acti
357551/1500000: episode: 236, duration: 14.437s, episode steps: 2360, steps per second: 163, episode reward: -11.000, mean reward: -0.005 [-1.000, 1.000], mean acti
360082/1500000: episode: 237, duration: 15.479s, episode steps: 2531, steps per second: 164, episode reward: -7.000, mean reward: -0.003 [-1.000, 1.000], mean acti
362739/1500000: episode: 238, duration: 15.984s, episode steps: 2657, steps per second: 166, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean acti
364881/1500000: episode: 239, duration: 12.850s, episode steps: 2142, steps per second: 167, episode reward: -13.000, mean reward: -0.006 [-1.000, 1.000], mean acti
367449/1500000: episode: 240, duration: 15.562s, episode steps: 2568, steps per second: 165, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean acti
370182/1500000: episode: 241, duration: 16.712s, episode steps: 2733, steps per second: 164, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
372345/1500000: episode: 242, duration: 13.158s, episode steps: 2163, steps per second: 164, episode reward: -9.000, mean reward: -0.004 [-1.000, 1.000], mean acti
374444/1500000: episode: 243, duration: 12.767s, episode steps: 2099, steps per second: 164, episode reward: -9.000, mean reward: -0.004 [-1.000, 1.000], mean acti
376617/1500000: episode: 244, duration: 13.327s, episode steps: 2173, steps per second: 163, episode reward: -13.000, mean reward: -0.006 [-1.000, 1.000], mean acti
378351/1500000: episode: 245, duration: 10.660s, episode steps: 1734, steps per second: 163, episode reward: -12.000, mean reward: -0.007 [-1.000, 1.000], mean acti
380210/1500000: episode: 246, duration: 11.451s, episode steps: 1859, steps per second: 162, episode reward: -13.000, mean reward: -0.007 [-1.000, 1.000], mean acti
382785/1500000: episode: 247, duration: 15.981s, episode steps: 2575, steps per second: 161, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean acti
385767/1500000: episode: 248, duration: 18.178s, episode steps: 2982, steps per second: 164, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
387973/1500000: episode: 249, duration: 13.468s, episode steps: 2206, steps per second: 164, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
390313/1500000: episode: 250, duration: 14.274s, episode steps: 2340, steps per second: 164, episode reward: -9.000, mean reward: -0.004 [-1.000, 1.000], mean acti
393325/1500000: episode: 251, duration: 18.717s, episode steps: 3012, steps per second: 161, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
395751/1500000: episode: 252, duration: 14.929s, episode steps: 2426, steps per second: 163, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
397286/1500000: episode: 253, duration: 9.422s, episode steps: 1535, steps per second: 163, episode reward: -14.000, mean reward: -0.009 [-1.000, 1.000], mean acti
399262/1500000: episode: 254, duration: 12.009s, episode steps: 1976, steps per second: 165, episode reward: -10.000, mean reward: -0.005 [-1.000, 1.000], mean acti
401963/1500000: episode: 255, duration: 16.476s, episode steps: 2701, steps per second: 164, episode reward: -2.000, mean reward: -0.001 [-1.000, 1.000], mean acti
402957/1500000: episode: 256, duration: 6.148s, episode steps: 994, steps per second: 162, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean actor
405811/1500000: episode: 257, duration: 17.396s, episode steps: 2854, steps per second: 164, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
408690/1500000: episode: 258, duration: 17.661s, episode steps: 2879, steps per second: 163, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
411590/1500000: episode: 259, duration: 17.731s, episode steps: 2900, steps per second: 164, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
413930/1500000: episode: 260, duration: 14.196s, episode steps: 2340, steps per second: 165, episode reward: -13.000, mean reward: -0.006 [-1.000, 1.000], mean acti
415517/1500000: episode: 261, duration: 9.700s, episode steps: 1587, steps per second: 164, episode reward: -17.000, mean reward: -0.011 [-1.000, 1.000], mean acti
418830/1500000: episode: 262, duration: 20.297s, episode steps: 3313, steps per second: 163, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
421616/1500000: episode: 263, duration: 17.220s, episode steps: 2786, steps per second: 162, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
424252/1500000: episode: 264, duration: 16.142s, episode steps: 2636, steps per second: 163, episode reward: -7.000, mean reward: -0.003 [-1.000, 1.000], mean acti
427138/1500000: episode: 265, duration: 17.614s, episode steps: 2866, steps per second: 164, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
430205/1500000: episode: 266, duration: 18.734s, episode steps: 3067, steps per second: 164, episode reward: -2.000, mean reward: -0.001 [-1.000, 1.000], mean acti
432108/1500000: episode: 267, duration: 11.695s, episode steps: 1903, steps per second: 163, episode reward: -13.000, mean reward: -0.007 [-1.000, 1.000], mean acti
434671/1500000: episode: 268, duration: 15.714s, episode steps: 2563, steps per second: 163, episode reward: -4.000, mean reward: -0.002 [-1.000, 1.000], mean acti
436744/1500000: episode: 269, duration: 12.511s, episode steps: 2073, steps per second: 166, episode reward: -11.000, mean reward: -0.005 [-1.000, 1.000], mean acti
439308/1500000: episode: 270, duration: 15.446s, episode steps: 2564, steps per second: 166, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean acti
442042/1500000: episode: 271, duration: 16.204s, episode steps: 2734, steps per second: 169, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean acti
443869/1500000: episode: 272, duration: 11.140s, episode steps: 1827, steps per second: 164, episode reward: -15.000, mean reward: -0.008 [-1.000, 1.000], mean acti
446783/1500000: episode: 273, duration: 17.662s, episode steps: 2914, steps per second: 165, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
449365/1500000: episode: 274, duration: 15.629s, episode steps: 2582, steps per second: 165, episode reward: -7.000, mean reward: -0.003 [-1.000, 1.000], mean acti
450943/1500000: episode: 275, duration: 9.456s, episode steps: 1578, steps per second: 167, episode reward: -18.000, mean reward: -0.011 [-1.000, 1.000], mean acti
453821/1500000: episode: 276, duration: 17.362s, episode steps: 2878, steps per second: 166, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
456453/1500000: episode: 277, duration: 15.799s, episode steps: 2632, steps per second: 167, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
458903/1500000: episode: 278, duration: 14.752s, episode steps: 2450, steps per second: 166, episode reward: 9.000, mean reward: 0.004 [-1.000, 1.000], mean action:
461603/1500000: episode: 279, duration: 16.306s, episode steps: 2700, steps per second: 166, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean acti
464081/1500000: episode: 280, duration: 14.845s, episode steps: 2478, steps per second: 167, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean acti
466229/1500000: episode: 281, duration: 12.801s, episode steps: 2148, steps per second: 168, episode reward: -8.000, mean reward: -0.004 [-1.000, 1.000], mean acti
468844/1500000: episode: 282, duration: 15.461s, episode steps: 2615, steps per second: 169, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean acti
470950/1500000: episode: 283, duration: 12.494s, episode steps: 2106, steps per second: 169, episode reward: -11.000, mean reward: -0.005 [-1.000, 1.000], mean acti
473421/1500000: episode: 284, duration: 14.617s, episode steps: 2471, steps per second: 169, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean acti
476233/1500000: episode: 285, duration: 16.795s, episode steps: 2812, steps per second: 167, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean acti
478741/1500000: episode: 286, duration: 14.982s, episode steps: 2508, steps per second: 167, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
481511/1500000: episode: 287, duration: 16.736s, episode steps: 2770, steps per second: 166, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
483722/1500000: episode: 288, duration: 13.250s, episode steps: 2211, steps per second: 167, episode reward: -8.000, mean reward: -0.004 [-1.000, 1.000], mean acti
486002/1500000: episode: 289, duration: 13.796s, episode steps: 2280, steps per second: 165, episode reward: -11.000, mean reward: -0.005 [-1.000, 1.000], mean acti
488782/1500000: episode: 290, duration: 16.832s, episode steps: 2780, steps per second: 165, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean acti
491181/1500000: episode: 291, duration: 14.675s, episode steps: 2399, steps per second: 163, episode reward: 7.000, mean reward: 0.003 [-1.000, 1.000], mean action:
493408/1500000: episode: 292, duration: 13.555s, episode steps: 2227, steps per second: 164, episode reward: -9.000, mean reward: -0.004 [-1.000, 1.000], mean acti
495506/1500000: episode: 293, duration: 13.113s, episode steps: 2098, steps per second: 160, episode reward: -11.000, mean reward: -0.005 [-1.000, 1.000], mean acti
498519/1500000: episode: 294, duration: 18.540s, episode steps: 3013, steps per second: 163, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
501460/1500000: episode: 295, duration: 18.066s, episode steps: 2941, steps per second: 163, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean acti
503723/1500000: episode: 296, duration: 13.900s, episode steps: 2263, steps per second: 163, episode reward: 13.000, mean reward: 0.006 [-1.000, 1.000], mean action:
506869/1500000: episode: 297, duration: 19.003s, episode steps: 3146, steps per second: 166, episode reward: -2.000, mean reward: -0.001 [-1.000, 1.000], mean acti
509428/1500000: episode: 298, duration: 15.624s, episode steps: 2559, steps per second: 164, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
512491/1500000: episode: 299, duration: 18.837s, episode steps: 3063, steps per second: 163, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean acti
515665/1500000: episode: 300, duration: 19.438s, episode steps: 3174, steps per second: 163, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
516886/1500000: episode: 301, duration: 7.289s, episode steps: 1221, steps per second: 168, episode reward: -19.000, mean reward: -0.016 [-1.000, 1.000], mean acti
519446/1500000: episode: 302, duration: 15.296s, episode steps: 2560, steps per second: 167, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean acti
521770/1500000: episode: 303, duration: 13.900s, episode steps: 2324, steps per second: 167, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
524228/1500000: episode: 304, duration: 14.651s, episode steps: 2458, steps per second: 168, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean acti
526834/1500000: episode: 305, duration: 15.677s, episode steps: 2606, steps per second: 166, episode reward: -7.000, mean reward: -0.003 [-1.000, 1.000], mean acti
529670/1500000: episode: 306, duration: 17.046s, episode steps: 2836, steps per second: 166, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
532298/1500000: episode: 307, duration: 15.810s, episode steps: 2628, steps per second: 166, episode reward: 7.000, mean reward: 0.003 [-1.000, 1.000], mean action:
535326/1500000: episode: 308, duration: 18.046s, episode steps: 3028, steps per second: 168, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
537708/1500000: episode: 309, duration: 14.288s, episode steps: 2382, steps per second: 167, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
540642/1500000: episode: 310, duration: 17.615s, episode steps: 2934, steps per second: 167, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
542910/1500000: episode: 311, duration: 13.618s, episode steps: 2268, steps per second: 167, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
545204/1500000: episode: 312, duration: 13.798s, episode steps: 2294, steps per second: 166, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
547544/1500000: episode: 313, duration: 14.254s, episode steps: 2340, steps per second: 164, episode reward: 13.000, mean reward: 0.006 [-1.000, 1.000], mean action:
550096/1500000: episode: 314, duration: 15.020s, episode steps: 2552, steps per second: 170, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
552698/1500000: episode: 315, duration: 15.673s, episode steps: 2602, steps per second: 166, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
555670/1500000: episode: 316, duration: 18.002s, episode steps: 2972, steps per second: 165, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
557918/1500000: episode: 317, duration: 13.480s, episode steps: 2248, steps per second: 167, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action:
561114/1500000: episode: 318, duration: 19.035s, episode steps: 3196, steps per second: 168, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
563510/1500000: episode: 319, duration: 14.257s, episode steps: 2396, steps per second: 168, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
565912/1500000: episode: 320, duration: 14.396s, episode steps: 2402, steps per second: 167, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean acti
568801/1500000: episode: 321, duration: 17.309s, episode steps: 2889, steps per second: 167, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
571638/1500000: episode: 322, duration: 16.709s, episode steps: 2837, steps per second: 170, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
574833/1500000: episode: 323, duration: 18.980s, episode steps: 3195, steps per second: 168, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
577744/1500000: episode: 324, duration: 17.476s, episode steps: 2911, steps per second: 167, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
580139/1500000: episode: 325, duration: 14.425s, episode steps: 2395, steps per second: 166, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
582916/1500000: episode: 326, duration: 16.630s, episode steps: 2777, steps per second: 167, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean acti

```

```

585541/1500000: episode: 327, duration: 15.752s, episode steps: 2625, steps per second: 167, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
588108/1500000: episode: 328, duration: 15.289s, episode steps: 2567, steps per second: 168, episode reward: 4.000, mean reward: 0.002 [-1.000, 1.000], mean action:
590373/1500000: episode: 329, duration: 13.540s, episode steps: 2265, steps per second: 167, episode reward: 9.000, mean reward: 0.004 [-1.000, 1.000], mean action:
592492/1500000: episode: 330, duration: 12.668s, episode steps: 2119, steps per second: 167, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action:
595404/1500000: episode: 331, duration: 17.461s, episode steps: 2912, steps per second: 167, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean action:
598194/1500000: episode: 332, duration: 16.922s, episode steps: 2790, steps per second: 165, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
601047/1500000: episode: 333, duration: 17.060s, episode steps: 2853, steps per second: 167, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
604164/1500000: episode: 334, duration: 18.550s, episode steps: 3117, steps per second: 168, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
606302/1500000: episode: 335, duration: 12.866s, episode steps: 2138, steps per second: 166, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action:
609118/1500000: episode: 336, duration: 16.740s, episode steps: 2816, steps per second: 168, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean action:
612320/1500000: episode: 337, duration: 19.107s, episode steps: 3202, steps per second: 168, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean action:
615023/1500000: episode: 338, duration: 16.072s, episode steps: 2703, steps per second: 168, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
617225/1500000: episode: 339, duration: 13.084s, episode steps: 2202, steps per second: 168, episode reward: -7.000, mean reward: -0.003 [-1.000, 1.000], mean action:
619618/1500000: episode: 340, duration: 14.164s, episode steps: 2393, steps per second: 169, episode reward: -4.000, mean reward: -0.002 [-1.000, 1.000], mean action:
621859/1500000: episode: 341, duration: 13.440s, episode steps: 2241, steps per second: 167, episode reward: 13.000, mean reward: 0.006 [-1.000, 1.000], mean action:
624675/1500000: episode: 342, duration: 16.699s, episode steps: 2816, steps per second: 169, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
626982/1500000: episode: 343, duration: 13.826s, episode steps: 2307, steps per second: 167, episode reward: 9.000, mean reward: 0.004 [-1.000, 1.000], mean action:
629552/1500000: episode: 344, duration: 15.325s, episode steps: 2570, steps per second: 168, episode reward: -4.000, mean reward: -0.002 [-1.000, 1.000], mean action:
632389/1500000: episode: 345, duration: 17.002s, episode steps: 2837, steps per second: 167, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
634748/1500000: episode: 346, duration: 14.120s, episode steps: 2359, steps per second: 167, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean action:
637229/1500000: episode: 347, duration: 14.850s, episode steps: 2481, steps per second: 167, episode reward: 9.000, mean reward: 0.004 [-1.000, 1.000], mean action:
640055/1500000: episode: 348, duration: 16.780s, episode steps: 2826, steps per second: 168, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
642727/1500000: episode: 349, duration: 15.854s, episode steps: 2672, steps per second: 169, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean action:
644922/1500000: episode: 350, duration: 12.936s, episode steps: 2195, steps per second: 170, episode reward: 11.000, mean reward: 0.005 [-1.000, 1.000], mean action:
647326/1500000: episode: 351, duration: 14.353s, episode steps: 2404, steps per second: 167, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
649734/1500000: episode: 352, duration: 14.582s, episode steps: 2408, steps per second: 165, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean action:
651793/1500000: episode: 353, duration: 12.443s, episode steps: 2059, steps per second: 165, episode reward: 16.000, mean reward: 0.008 [-1.000, 1.000], mean action:
654574/1500000: episode: 354, duration: 16.569s, episode steps: 2781, steps per second: 168, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean action:
656781/1500000: episode: 355, duration: 13.301s, episode steps: 2207, steps per second: 166, episode reward: -11.000, mean reward: -0.005 [-1.000, 1.000], mean action:
659608/1500000: episode: 356, duration: 16.863s, episode steps: 2827, steps per second: 168, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean action:
662493/1500000: episode: 357, duration: 17.375s, episode steps: 2885, steps per second: 166, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
665008/1500000: episode: 358, duration: 15.048s, episode steps: 2515, steps per second: 167, episode reward: 9.000, mean reward: 0.004 [-1.000, 1.000], mean action:
668358/1500000: episode: 359, duration: 20.114s, episode steps: 3350, steps per second: 167, episode reward: -2.000, mean reward: -0.001 [-1.000, 1.000], mean action:
671593/1500000: episode: 360, duration: 19.285s, episode steps: 3235, steps per second: 168, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
674081/1500000: episode: 361, duration: 14.902s, episode steps: 2488, steps per second: 167, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
675805/1500000: episode: 362, duration: 10.331s, episode steps: 1724, steps per second: 167, episode reward: 20.000, mean reward: 0.012 [-1.000, 1.000], mean action:
678444/1500000: episode: 363, duration: 15.892s, episode steps: 2639, steps per second: 166, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
680795/1500000: episode: 364, duration: 14.099s, episode steps: 2351, steps per second: 167, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
684112/1500000: episode: 365, duration: 19.720s, episode steps: 3317, steps per second: 168, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
686759/1500000: episode: 366, duration: 15.844s, episode steps: 2647, steps per second: 167, episode reward: 7.000, mean reward: 0.003 [-1.000, 1.000], mean action:
688824/1500000: episode: 367, duration: 12.485s, episode steps: 2065, steps per second: 165, episode reward: 13.000, mean reward: 0.006 [-1.000, 1.000], mean action:
691698/1500000: episode: 368, duration: 17.196s, episode steps: 2874, steps per second: 167, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
694477/1500000: episode: 369, duration: 16.597s, episode steps: 2779, steps per second: 167, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean action:
697090/1500000: episode: 370, duration: 15.551s, episode steps: 2613, steps per second: 168, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean action:
699073/1500000: episode: 371, duration: 11.873s, episode steps: 1983, steps per second: 167, episode reward: -13.000, mean reward: -0.007 [-1.000, 1.000], mean action:
701263/1500000: episode: 372, duration: 13.321s, episode steps: 2190, steps per second: 164, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
704194/1500000: episode: 373, duration: 17.599s, episode steps: 2931, steps per second: 167, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
706171/1500000: episode: 374, duration: 11.886s, episode steps: 1977, steps per second: 166, episode reward: 16.000, mean reward: 0.008 [-1.000, 1.000], mean action:
708326/1500000: episode: 375, duration: 12.827s, episode steps: 2155, steps per second: 168, episode reward: 12.000, mean reward: 0.006 [-1.000, 1.000], mean action:
710319/1500000: episode: 376, duration: 12.035s, episode steps: 1993, steps per second: 166, episode reward: 14.000, mean reward: 0.007 [-1.000, 1.000], mean action:
712780/1500000: episode: 377, duration: 14.738s, episode steps: 2461, steps per second: 167, episode reward: 7.000, mean reward: 0.003 [-1.000, 1.000], mean action:
715111/1500000: episode: 378, duration: 13.999s, episode steps: 2331, steps per second: 167, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
717834/1500000: episode: 379, duration: 16.180s, episode steps: 2723, steps per second: 168, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
720066/1500000: episode: 380, duration: 13.390s, episode steps: 2232, steps per second: 167, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
722683/1500000: episode: 381, duration: 15.633s, episode steps: 2617, steps per second: 167, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
725699/1500000: episode: 382, duration: 17.923s, episode steps: 3016, steps per second: 168, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean action:
727939/1500000: episode: 383, duration: 13.447s, episode steps: 2240, steps per second: 167, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
730589/1500000: episode: 384, duration: 15.841s, episode steps: 2650, steps per second: 167, episode reward: -4.000, mean reward: -0.002 [-1.000, 1.000], mean action:
733153/1500000: episode: 385, duration: 15.314s, episode steps: 2564, steps per second: 167, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
736026/1500000: episode: 386, duration: 17.028s, episode steps: 2873, steps per second: 169, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
738484/1500000: episode: 387, duration: 14.591s, episode steps: 2458, steps per second: 168, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
740546/1500000: episode: 388, duration: 12.312s, episode steps: 2062, steps per second: 167, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action:
743446/1500000: episode: 389, duration: 17.826s, episode steps: 2900, steps per second: 163, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
746432/1500000: episode: 390, duration: 17.879s, episode steps: 2986, steps per second: 167, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
748951/1500000: episode: 391, duration: 15.030s, episode steps: 2519, steps per second: 168, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
751309/1500000: episode: 392, duration: 14.154s, episode steps: 2358, steps per second: 167, episode reward: 9.000, mean reward: 0.004 [-1.000, 1.000], mean action:
753549/1500000: episode: 393, duration: 13.722s, episode steps: 2240, steps per second: 163, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
756366/1500000: episode: 394, duration: 16.972s, episode steps: 2817, steps per second: 166, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
759005/1500000: episode: 395, duration: 15.838s, episode steps: 2639, steps per second: 167, episode reward: 7.000, mean reward: 0.003 [-1.000, 1.000], mean action:
761746/1500000: episode: 396, duration: 16.299s, episode steps: 2741, steps per second: 168, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean action:
764720/1500000: episode: 397, duration: 17.674s, episode steps: 2974, steps per second: 168, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
766951/1500000: episode: 398, duration: 13.300s, episode steps: 2231, steps per second: 168, episode reward: 13.000, mean reward: 0.006 [-1.000, 1.000], mean action:
769609/1500000: episode: 399, duration: 15.815s, episode steps: 2658, steps per second: 168, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean action:
771883/1500000: episode: 400, duration: 13.437s, episode steps: 2274, steps per second: 169, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action:
774510/1500000: episode: 401, duration: 15.612s, episode steps: 2627, steps per second: 168, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
776680/1500000: episode: 402, duration: 12.848s, episode steps: 2170, steps per second: 169, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean action:
779050/1500000: episode: 403, duration: 14.186s, episode steps: 2370, steps per second: 167, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action:
780694/1500000: episode: 404, duration: 9.855s, episode steps: 1644, steps per second: 167, episode reward: 21.000, mean reward: 0.013 [0.000, 1.000], mean action:
782338/1500000: episode: 405, duration: 9.982s, episode steps: 1644, steps per second: 165, episode reward: 21.000, mean reward: 0.013 [0.000, 1.000], mean action:
785166/1500000: episode: 406, duration: 17.080s, episode steps: 2828, steps per second: 166, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
787757/1500000: episode: 407, duration: 15.802s, episode steps: 2591, steps per second: 164, episode reward: -2.000, mean reward: -0.001 [-1.000, 1.000], mean action:
790025/1500000: episode: 408, duration: 13.859s, episode steps: 2268, steps per second: 164, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action:
792849/1500000: episode: 409, duration: 16.920s, episode steps: 2824, steps per second: 167, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean action:
795272/1500000: episode: 410, duration: 14.886s, episode steps: 2423, steps per second: 163, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
797478/1500000: episode: 411, duration: 13.286s, episode steps: 2206, steps per second: 166, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action:
799401/1500000: episode: 412, duration: 11.679s, episode steps: 1923, steps per second: 165, episode reward: 17.000, mean reward: 0.009 [-1.000, 1.000], mean action:
801238/1500000: episode: 413, duration: 11.080s, episode steps: 1837, steps per second: 166, episode reward: 16.000, mean reward: 0.009 [-1.000, 1.000], mean action:
803331/1500000: episode: 414, duration: 12.825s, episode steps: 2093, steps per second: 163, episode reward: 13.000, mean reward: 0.006 [-1.000, 1.000], mean action:
805619/1500000: episode: 415, duration: 14.293s, episode steps: 2288, steps per second: 160, episode reward: 13.000, mean reward: 0.006 [-1.000, 1.000], mean action:
808169/1500000: episode: 416, duration: 15.664s, episode steps: 2550, steps per second: 163, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
810987/1500000: episode: 417, duration: 17.017s, episode steps: 2818, steps per second: 166, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
813715/1500000: episode: 418, duration: 16.436s, episode steps: 2728, steps per second: 166, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
815735/1500000: episode: 419, duration: 12.147s, episode steps: 2020, steps per second: 166, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action:
817791/1500000: episode: 420, duration: 12.275s, episode steps: 2056, steps per second: 168, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action:
820447/1500000: episode: 421, duration: 16.192s, episode steps: 2656, steps per second: 164, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean action:
822861/1500000: episode: 422, duration: 14.485s, episode steps: 2414, steps per second: 167, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action:
825153/1500000: episode: 423, duration: 13.677s, episode steps: 2292, steps per second: 168, episode reward: 11.000, mean reward: 0.005 [-1.000, 1.000], mean action:
827592/1500000: episode: 424, duration: 14.531s, episode steps: 2439, steps per second: 168, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
829967/1500000: episode: 425, duration: 14.245s, episode steps: 2375, steps per second: 167, episode reward: 11.000, mean reward: 0.005 [-1.000, 1.000], mean action:
832319/1500000: episode: 426, duration: 14.094s, episode steps: 2352, steps per second: 167, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
834416/1500000: episode: 427, duration: 12.714s, episode steps: 2097, steps per second: 165, episode reward: 14.000, mean reward: 0.007 [-1.000, 1.000], mean action:
836553/1500000: episode: 428, duration: 12.747s, episode steps: 2137, steps per second: 168, episode reward: 14.000, mean reward: 0.007 [-1.000, 1.000], mean action:
839104/1500000: episode: 429, duration: 15.236s, episode steps: 2551, steps per second: 167, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
841188/1500000: episode: 430, duration: 12.445s, episode steps: 2084, steps per second: 167, episode reward: 13.000, mean reward: 0.006 [-1.000, 1.000], mean action:
843178/1500000: episode: 431, duration: 11.901s, episode steps: 1990, steps per second: 167, episode reward: 15.000, mean reward: 0.008 [-1.000, 1.000], mean action:
845063/1500000: episode: 432, duration: 11.318s, episode steps: 1885, steps per second: 167, episode reward: 17.000, mean reward: 0.009 [-1.000, 1.000], mean action:
847883/1500000: episode: 433, duration: 16.962s, episode steps: 2820, steps per second: 166, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
850801/1500000: episode: 434, duration: 17.463s, episode steps: 2918, steps per second: 167, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
853108/1500000: episode: 435, duration: 13.736s, episode steps: 2307, steps per second: 168, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:

```

855805/1500000:	episode:	436,	duration:	16.331s,	episode steps:	2697,	steps per second:	165,	episode reward:	7.000,	mean reward:	0.003 [-1.000, 1.000],	mean action:
858157/1500000:	episode:	437,	duration:	14.268s,	episode steps:	2352,	steps per second:	165,	episode reward:	10.000,	mean reward:	0.004 [-1.000, 1.000],	mean action:
860319/1500000:	episode:	438,	duration:	12.677s,	episode steps:	2162,	steps per second:	171,	episode reward:	13.000,	mean reward:	0.006 [-1.000, 1.000],	mean action:
863137/1500000:	episode:	439,	duration:	16.922s,	episode steps:	2818,	steps per second:	167,	episode reward:	3.000,	mean reward:	0.001 [-1.000, 1.000],	mean action:
865706/1500000:	episode:	440,	duration:	15.358s,	episode steps:	2569,	steps per second:	167,	episode reward:	6.000,	mean reward:	0.002 [-1.000, 1.000],	mean action:
868695/1500000:	episode:	441,	duration:	17.902s,	episode steps:	2989,	steps per second:	167,	episode reward:	5.000,	mean reward:	0.002 [-1.000, 1.000],	mean action:
871411/1500000:	episode:	442,	duration:	16.196s,	episode steps:	2716,	steps per second:	168,	episode reward:	8.000,	mean reward:	0.003 [-1.000, 1.000],	mean action:
873405/1500000:	episode:	443,	duration:	11.906s,	episode steps:	1994,	steps per second:	167,	episode reward:	15.000,	mean reward:	0.008 [-1.000, 1.000],	mean action:
875798/1500000:	episode:	444,	duration:	14.278s,	episode steps:	2393,	steps per second:	168,	episode reward:	7.000,	mean reward:	0.003 [-1.000, 1.000],	mean action:
877752/1500000:	episode:	445,	duration:	11.815s,	episode steps:	1954,	steps per second:	165,	episode reward:	-12.000,	mean reward:	-0.006 [-1.000, 1.000],	mean action:
880245/1500000:	episode:	446,	duration:	14.846s,	episode steps:	2493,	steps per second:	168,	episode reward:	11.000,	mean reward:	0.004 [-1.000, 1.000],	mean action:
883031/1500000:	episode:	447,	duration:	16.678s,	episode steps:	2786,	steps per second:	167,	episode reward:	-2.000,	mean reward:	-0.001 [-1.000, 1.000],	mean action:
885386/1500000:	episode:	448,	duration:	14.059s,	episode steps:	2355,	steps per second:	168,	episode reward:	12.000,	mean reward:	0.005 [-1.000, 1.000],	mean action:
888087/1500000:	episode:	449,	duration:	16.163s,	episode steps:	2701,	steps per second:	167,	episode reward:	6.000,	mean reward:	0.002 [-1.000, 1.000],	mean action:
890768/1500000:	episode:	450,	duration:	16.001s,	episode steps:	2681,	steps per second:	168,	episode reward:	5.000,	mean reward:	0.002 [-1.000, 1.000],	mean action:
893126/1500000:	episode:	451,	duration:	14.030s,	episode steps:	2358,	steps per second:	168,	episode reward:	15.000,	mean reward:	0.006 [-1.000, 1.000],	mean action:
895444/1500000:	episode:	452,	duration:	13.816s,	episode steps:	2318,	steps per second:	168,	episode reward:	11.000,	mean reward:	0.005 [-1.000, 1.000],	mean action:
898242/1500000:	episode:	453,	duration:	16.764s,	episode steps:	2798,	steps per second:	167,	episode reward:	-4.000,	mean reward:	-0.001 [-1.000, 1.000],	mean action:
900420/1500000:	episode:	454,	duration:	12.832s,	episode steps:	2178,	steps per second:	170,	episode reward:	13.000,	mean reward:	0.006 [-1.000, 1.000],	mean action:
903215/1500000:	episode:	455,	duration:	16.495s,	episode steps:	2795,	steps per second:	169,	episode reward:	4.000,	mean reward:	0.001 [-1.000, 1.000],	mean action:
905235/1500000:	episode:	456,	duration:	12.084s,	episode steps:	2020,	steps per second:	167,	episode reward:	15.000,	mean reward:	0.007 [-1.000, 1.000],	mean action:
908339/1500000:	episode:	457,	duration:	19.062s,	episode steps:	3104,	steps per second:	163,	episode reward:	1.000,	mean reward:	0.000 [-1.000, 1.000],	mean action:
910562/1500000:	episode:	458,	duration:	13.367s,	episode steps:	2223,	steps per second:	166,	episode reward:	13.000,	mean reward:	0.006 [-1.000, 1.000],	mean action:
912860/1500000:	episode:	459,	duration:	13.726s,	episode steps:	2298,	steps per second:	167,	episode reward:	9.000,	mean reward:	0.004 [-1.000, 1.000],	mean action:
915356/1500000:	episode:	460,	duration:	14.943s,	episode steps:	2496,	steps per second:	167,	episode reward:	8.000,	mean reward:	0.003 [-1.000, 1.000],	mean action:
917551/1500000:	episode:	461,	duration:	13.090s,	episode steps:	2195,	steps per second:	168,	episode reward:	12.000,	mean reward:	0.005 [-1.000, 1.000],	mean action:
919869/1500000:	episode:	462,	duration:	13.847s,	episode steps:	2318,	steps per second:	167,	episode reward:	10.000,	mean reward:	0.004 [-1.000, 1.000],	mean action:
922613/1500000:	episode:	463,	duration:	16.243s,	episode steps:	2744,	steps per second:	169,	episode reward:	7.000,	mean reward:	0.003 [-1.000, 1.000],	mean action:
924903/1500000:	episode:	464,	duration:	13.625s,	episode steps:	2290,	steps per second:	168,	episode reward:	10.000,	mean reward:	0.004 [-1.000, 1.000],	mean action:
926784/1500000:	episode:	465,	duration:	11.203s,	episode steps:	1881,	steps per second:	168,	episode reward:	17.000,	mean reward:	0.009 [-1.000, 1.000],	mean action:
929063/1500000:	episode:	466,	duration:	13.573s,	episode steps:	2279,	steps per second:	168,	episode reward:	13.000,	mean reward:	0.006 [-1.000, 1.000],	mean action:
931032/1500000:	episode:	467,	duration:	11.737s,	episode steps:	1969,	steps per second:	168,	episode reward:	18.000,	mean reward:	0.009 [-1.000, 1.000],	mean action:
933673/1500000:	episode:	468,	duration:	15.721s,	episode steps:	2641,	steps per second:	168,	episode reward:	-1.000,	mean reward:	-0.000 [-1.000, 1.000],	mean action:
936374/1500000:	episode:	469,	duration:	15.992s,	episode steps:	2701,	steps per second:	169,	episode reward:	5.000,	mean reward:	0.002 [-1.000, 1.000],	mean action:
938901/1500000:	episode:	470,	duration:	14.978s,	episode steps:	2527,	steps per second:	169,	episode reward:	7.000,	mean reward:	0.003 [-1.000, 1.000],	mean action:
941424/1500000:	episode:	471,	duration:	14.945s,	episode steps:	2523,	steps per second:	169,	episode reward:	7.000,	mean reward:	0.003 [-1.000, 1.000],	mean action:
943979/1500000:	episode:	472,	duration:	14.852s,	episode steps:	2555,	steps per second:	172,	episode reward:	7.000,	mean reward:	0.003 [-1.000, 1.000],	mean action:
946138/1500000:	episode:	473,	duration:	12.640s,	episode steps:	2159,	steps per second:	171,	episode reward:	11.000,	mean reward:	0.005 [-1.000, 1.000],	mean action:
949356/1500000:	episode:	474,	duration:	18.866s,	episode steps:	3218,	steps per second:	171,	episode reward:	1.000,	mean reward:	0.000 [-1.000, 1.000],	mean action:
951626/1500000:	episode:	475,	duration:	13.477s,	episode steps:	2270,	steps per second:	168,	episode reward:	11.000,	mean reward:	0.005 [-1.000, 1.000],	mean action:
954256/1500000:	episode:	476,	duration:	15.520s,	episode steps:	2630,	steps per second:	169,	episode reward:	5.000,	mean reward:	0.002 [-1.000, 1.000],	mean action:
957361/1500000:	episode:	477,	duration:	18.410s,	episode steps:	3105,	steps per second:	169,	episode reward:	1.000,	mean reward:	0.000 [-1.000, 1.000],	mean action:
960552/1500000:	episode:	478,	duration:	19.270s,	episode steps:	3191,	steps per second:	166,	episode reward:	1.000,	mean reward:	0.000 [-1.000, 1.000],	mean action:
963627/1500000:	episode:	479,	duration:	18.199s,	episode steps:	3075,	steps per second:	169,	episode reward:	2.000,	mean reward:	0.001 [-1.000, 1.000],	mean action:
966257/1500000:	episode:	480,	duration:	15.561s,	episode steps:	2630,	steps per second:	169,	episode reward:	-9.000,	mean reward:	-0.003 [-1.000, 1.000],	mean action:
969447/1500000:	episode:	481,	duration:	18.914s,	episode steps:	3190,	steps per second:	169,	episode reward:	-3.000,	mean reward:	-0.001 [-1.000, 1.000],	mean action:
972024/1500000:	episode:	482,	duration:	15.313s,	episode steps:	2577,	steps per second:	168,	episode reward:	14.000,	mean reward:	0.005 [-1.000, 1.000],	mean action:
974233/1500000:	episode:	483,	duration:	13.134s,	episode steps:	2209,	steps per second:	168,	episode reward:	15.000,	mean reward:	0.007 [-1.000, 1.000],	mean action:
976726/1500000:	episode:	484,	duration:	14.796s,	episode steps:	2493,	steps per second:	168,	episode reward:	12.000,	mean reward:	0.005 [-1.000, 1.000],	mean action:
978869/1500000:	episode:	485,	duration:	12.712s,	episode steps:	2143,	steps per second:	169,	episode reward:	14.000,	mean reward:	0.007 [-1.000, 1.000],	mean action:
981351/1500000:	episode:	486,	duration:	14.764s,	episode steps:	2482,	steps per second:	168,	episode reward:	6.000,	mean reward:	0.002 [-1.000, 1.000],	mean action:
983527/1500000:	episode:	487,	duration:	12.975s,	episode steps:	2176,	steps per second:	168,	episode reward:	15.000,	mean reward:	0.007 [-1.000, 1.000],	mean action:
985590/1500000:	episode:	488,	duration:	12.310s,	episode steps:	2063,	steps per second:	168,	episode reward:	15.000,	mean reward:	0.007 [-1.000, 1.000],	mean action:
987563/1500000:	episode:	489,	duration:	11.738s,	episode steps:	1973,	steps per second:	168,	episode reward:	17.000,	mean reward:	0.009 [-1.000, 1.000],	mean action:
990410/1500000:	episode:	490,	duration:	16.778s,	episode steps:	2847,	steps per second:	170,	episode reward:	4.000,	mean reward:	0.001 [-1.000, 1.000],	mean action:
992581/1500000:	episode:	491,	duration:	12.697s,	episode steps:	2171,	steps per second:	171,	episode reward:	16.000,	mean reward:	0.007 [-1.000, 1.000],	mean action:
994524/1500000:	episode:	492,	duration:	11.678s,	episode steps:	1943,	steps per second:	166,	episode reward:	17.000,	mean reward:	0.009 [-1.000, 1.000],	mean action:
996571/1500000:	episode:	493,	duration:	12.260s,	episode steps:	2047,	steps per second:	167,	episode reward:	16.000,	mean reward:	0.008 [-1.000, 1.000],	mean action:
998836/1500000:	episode:	494,	duration:	13.509s,	episode steps:	2265,	steps per second:	168,	episode reward:	11.000,	mean reward:	0.005 [-1.000, 1.000],	mean action:
1001118/1500000:	episode:	495,	duration:	13.609s,	episode steps:	2282,	steps per second:	168,	episode reward:	13.000,	mean reward:	0.006 [-1.000, 1.000],	mean action:
1003234/1500000:	episode:	496,	duration:	12.653s,	episode steps:	2116,	steps per second:	167,	episode reward:	13.000,	mean reward:	0.006 [-1.000, 1.000],	mean action:
1005661/1500000:	episode:	497,	duration:	14.501s,	episode steps:	2427,	steps per second:	167,	episode reward:	11.000,	mean reward:	0.005 [-1.000, 1.000],	mean action:
1008026/1500000:	episode:	498,	duration:	14.032s,	episode steps:	2365,	steps per second:	169,	episode reward:	8.000,	mean reward:	0.003 [-1.000, 1.000],	mean action:
1011056/1500000:	episode:	499,	duration:	18.376s,	episode steps:	3030,	steps per second:	165,	episode reward:	-1.000,	mean reward:	-0.000 [-1.000, 1.000],	mean action:
1013595/1500000:	episode:	500,	duration:	15.335s,	episode steps:	2539,	steps per second:	166,	episode reward:	8.000,	mean reward:	0.003 [-1.000, 1.000],	mean action:
1016042/1500000:	episode:	501,	duration:	14.538s,	episode steps:	2447,	steps per second:	168,	episode reward:	10.000,	mean reward:	0.004 [-1.000, 1.000],	mean action:
1018956/1500000:	episode:	502,	duration:	17.293s,	episode steps:	2914,	steps per second:	169,	episode reward:	3.000,	mean reward:	0.001 [-1.000, 1.000],	mean action:
1021386/1500000:	episode:	503,	duration:	14.520s,	episode steps:	2430,	steps per second:	167,	episode reward:	12.000,	mean reward:	0.005 [-1.000, 1.000],	mean action:
1024061/1500000:	episode:	504,	duration:	16.000s,	episode steps:	2675,	steps per second:	167,	episode reward:	4.000,	mean reward:	0.001 [-1.000, 1.000],	mean action:
1025986/1500000:	episode:	505,	duration:	11.454s,	episode steps:	1925,	steps per second:	168,	episode reward:	17.000,	mean reward:	0.009 [-1.000, 1.000],	mean action:
1028329/1500000:	episode:	506,	duration:	13.941s,	episode steps:	2343,	steps per second:	168,	episode reward:	12.000,	mean reward:	0.005 [-1.000, 1.000],	mean action:
1030994/1500000:	episode:	507,	duration:	15.858s,	episode steps:	2665,	steps per second:	168,	episode reward:	9.000,	mean reward:	0.003 [-1.000, 1.000],	mean action:
1033610/1500000:	episode:	508,	duration:	15.400s,	episode steps:	2616,	steps per second:	170,	episode reward:	5.000,	mean reward:	0.002 [-1.000, 1.000],	mean action:
1035601/1500000:	episode:	509,	duration:	11.997s,	episode steps:	1991,	steps per second:	166,	episode reward:	16.000,	mean reward:	0.008 [-1.000, 1.000],	mean action:
1038652/1500000:	episode:	510,	duration:	18.187s,	episode steps:	3051,	steps per second:	168,	episode reward:	1.000,	mean reward:	0.000 [-1.000, 1.000],	mean action:
1040427/1500000:	episode:	511,	duration:	10.560s,	episode steps:	1775,	steps per second:	168,	episode reward:	19.000,	mean reward:	0.011 [-1.000, 1.000],	mean action:
1042718/1500000:	episode:	512,	duration:	13.631s,	episode steps:	2291,	steps per second:	168,	episode reward:	15.000,	mean reward:	0.007 [-1.000, 1.000],	mean action:
1044854/1500000:	episode:	513,	duration:	12.750s,	episode steps:	2136,	steps per second:	168,	episode reward:	16.000,	mean reward:	0.007 [-1.000, 1.000],	mean action:
1047981/1500000:	episode:	514,	duration:	18.641s,	episode steps:	3127,	steps per second:	168,	episode reward:	5.000,	mean reward:	0.002 [-1.000, 1.000],	mean action:
1051308/1500000:	episode:	515,	duration:	19.943s,	episode steps:	3327,	steps per second:	167,	episode reward:	-2.000,	mean reward:	-0.001 [-1.000, 1.000],	mean action:
1054558/1500000:	episode:	516,	duration:	19.323s,	episode steps:	3250,	steps per second:	168,	episode reward:	2.000,	mean reward:	0.001 [-1.000, 1.000],	mean action:
1057353/1500000:	episode:	517,	duration:	16.702s,	episode steps:	2795,	steps per second:	167,	episode reward:	1.000,	mean reward:	0.000 [-1.000, 1.000],	mean action:
1060347/1500000:	episode:	518,	duration:	17.860s,	episode steps:	2994,	steps per second:	168,	episode reward:	7.000,	mean reward:	0.002 [-1.000, 1.000],	mean action:
1062812/1500000:	episode:	519,	duration:	14.992s,	episode steps:	2465,	steps per second:	164,	episode reward:	14.000,	mean reward:	0.006 [-1.000, 1.000],	mean action:
1065636/1500000:	episode:	520,	duration:	17.129s,	episode steps:	2824,	steps per second:	165,	episode reward:	7.000,	mean reward:	0.002 [-1.000, 1.000],	mean action:
1068837/1500000:	episode:	521,	duration:	19.199s,	episode steps:								

```

1130463/1500000: episode: 545, duration: 16.603s, episode steps: 2745, steps per second: 165, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1133007/1500000: episode: 546, duration: 15.259s, episode steps: 2544, steps per second: 167, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1134849/1500000: episode: 547, duration: 11.118s, episode steps: 1842, steps per second: 166, episode reward: 16.000, mean reward: 0.009 [-1.000, 1.000], mean action:
1137253/1500000: episode: 548, duration: 14.447s, episode steps: 2404, steps per second: 166, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1139327/1500000: episode: 549, duration: 12.476s, episode steps: 2074, steps per second: 166, episode reward: 16.000, mean reward: 0.008 [-1.000, 1.000], mean action:
1141997/1500000: episode: 550, duration: 16.028s, episode steps: 2670, steps per second: 167, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
1143921/1500000: episode: 551, duration: 11.580s, episode steps: 1924, steps per second: 166, episode reward: 17.000, mean reward: 0.009 [-1.000, 1.000], mean action:
1146386/1500000: episode: 552, duration: 14.750s, episode steps: 2465, steps per second: 167, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1148807/1500000: episode: 553, duration: 14.541s, episode steps: 2421, steps per second: 166, episode reward: 11.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1151324/1500000: episode: 554, duration: 15.149s, episode steps: 2517, steps per second: 166, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1153288/1500000: episode: 555, duration: 11.732s, episode steps: 1964, steps per second: 167, episode reward: 15.000, mean reward: 0.008 [-1.000, 1.000], mean action:
1155351/1500000: episode: 556, duration: 12.504s, episode steps: 2063, steps per second: 165, episode reward: 16.000, mean reward: 0.008 [-1.000, 1.000], mean action:
1157826/1500000: episode: 557, duration: 14.861s, episode steps: 2475, steps per second: 167, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1160040/1500000: episode: 558, duration: 13.395s, episode steps: 2214, steps per second: 165, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1163184/1500000: episode: 559, duration: 18.909s, episode steps: 3144, steps per second: 166, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean action:
1165223/1500000: episode: 560, duration: 12.325s, episode steps: 2039, steps per second: 165, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action:
1167025/1500000: episode: 561, duration: 10.975s, episode steps: 1802, steps per second: 164, episode reward: -10.000, mean reward: -0.006 [-1.000, 1.000], mean action:
1169498/1500000: episode: 562, duration: 14.943s, episode steps: 2473, steps per second: 165, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1171948/1500000: episode: 563, duration: 14.715s, episode steps: 2450, steps per second: 166, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1174443/1500000: episode: 564, duration: 14.873s, episode steps: 2495, steps per second: 168, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1176812/1500000: episode: 565, duration: 14.314s, episode steps: 2369, steps per second: 166, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1179225/1500000: episode: 566, duration: 14.593s, episode steps: 2413, steps per second: 165, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1181906/1500000: episode: 567, duration: 16.238s, episode steps: 2681, steps per second: 165, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean action:
1183724/1500000: episode: 568, duration: 10.950s, episode steps: 1818, steps per second: 166, episode reward: 20.000, mean reward: 0.011 [-1.000, 1.000], mean action:
1185761/1500000: episode: 569, duration: 12.339s, episode steps: 2037, steps per second: 165, episode reward: 14.000, mean reward: 0.007 [-1.000, 1.000], mean action:
1188818/1500000: episode: 570, duration: 18.316s, episode steps: 3057, steps per second: 167, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1191834/1500000: episode: 571, duration: 18.136s, episode steps: 3016, steps per second: 166, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1193550/1500000: episode: 572, duration: 10.246s, episode steps: 1716, steps per second: 167, episode reward: 20.000, mean reward: 0.012 [-1.000, 1.000], mean action:
1195707/1500000: episode: 573, duration: 12.937s, episode steps: 2157, steps per second: 167, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action:
1198397/1500000: episode: 574, duration: 15.930s, episode steps: 2672, steps per second: 168, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1200666/1500000: episode: 575, duration: 13.574s, episode steps: 2287, steps per second: 168, episode reward: 13.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1203040/1500000: episode: 576, duration: 14.330s, episode steps: 2374, steps per second: 166, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1205301/1500000: episode: 577, duration: 13.502s, episode steps: 2261, steps per second: 167, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean action:
1207587/1500000: episode: 578, duration: 13.588s, episode steps: 2286, steps per second: 168, episode reward: 11.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1210969/1500000: episode: 579, duration: 20.088s, episode steps: 3382, steps per second: 168, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1214186/1500000: episode: 580, duration: 19.420s, episode steps: 3217, steps per second: 166, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1216804/1500000: episode: 581, duration: 15.746s, episode steps: 2618, steps per second: 166, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1219654/1500000: episode: 582, duration: 17.155s, episode steps: 2850, steps per second: 166, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1222154/1500000: episode: 583, duration: 14.946s, episode steps: 2500, steps per second: 167, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1224483/1500000: episode: 584, duration: 13.906s, episode steps: 2329, steps per second: 167, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1226886/1500000: episode: 585, duration: 14.282s, episode steps: 2403, steps per second: 168, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1228941/1500000: episode: 586, duration: 12.217s, episode steps: 2055, steps per second: 168, episode reward: 16.000, mean reward: 0.008 [-1.000, 1.000], mean action:
1231630/1500000: episode: 587, duration: 15.984s, episode steps: 2689, steps per second: 168, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1233641/1500000: episode: 588, duration: 12.072s, episode steps: 2011, steps per second: 167, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean action:
1236864/1500000: episode: 589, duration: 19.294s, episode steps: 3223, steps per second: 167, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1238820/1500000: episode: 590, duration: 11.651s, episode steps: 1956, steps per second: 168, episode reward: 19.000, mean reward: 0.010 [-1.000, 1.000], mean action:
1241415/1500000: episode: 591, duration: 15.373s, episode steps: 2595, steps per second: 169, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1243897/1500000: episode: 592, duration: 14.802s, episode steps: 2482, steps per second: 168, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1246109/1500000: episode: 593, duration: 13.065s, episode steps: 2212, steps per second: 169, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1249142/1500000: episode: 594, duration: 18.296s, episode steps: 3033, steps per second: 166, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1252677/1500000: episode: 595, duration: 20.950s, episode steps: 3535, steps per second: 169, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1255146/1500000: episode: 596, duration: 14.805s, episode steps: 2469, steps per second: 167, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1257953/1500000: episode: 597, duration: 16.783s, episode steps: 2807, steps per second: 167, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1261057/1500000: episode: 598, duration: 18.723s, episode steps: 3104, steps per second: 166, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1263227/1500000: episode: 599, duration: 13.187s, episode steps: 2170, steps per second: 165, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean action:
1265738/1500000: episode: 600, duration: 15.158s, episode steps: 2511, steps per second: 166, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1268636/1500000: episode: 601, duration: 17.634s, episode steps: 2898, steps per second: 164, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1271407/1500000: episode: 602, duration: 16.826s, episode steps: 2771, steps per second: 165, episode reward: 7.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1274414/1500000: episode: 603, duration: 18.139s, episode steps: 3007, steps per second: 166, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1276427/1500000: episode: 604, duration: 12.186s, episode steps: 2013, steps per second: 165, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean action:
1278710/1500000: episode: 605, duration: 13.802s, episode steps: 2283, steps per second: 165, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action:
1281615/1500000: episode: 606, duration: 17.534s, episode steps: 2905, steps per second: 166, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1283448/1500000: episode: 607, duration: 11.083s, episode steps: 1833, steps per second: 165, episode reward: 17.000, mean reward: 0.009 [-1.000, 1.000], mean action:
1286215/1500000: episode: 608, duration: 16.746s, episode steps: 2767, steps per second: 165, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1288574/1500000: episode: 609, duration: 14.250s, episode steps: 2359, steps per second: 166, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1290954/1500000: episode: 610, duration: 14.343s, episode steps: 2380, steps per second: 166, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1294003/1500000: episode: 611, duration: 18.248s, episode steps: 3049, steps per second: 167, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1295991/1500000: episode: 612, duration: 11.851s, episode steps: 1988, steps per second: 168, episode reward: 17.000, mean reward: 0.009 [-1.000, 1.000], mean action:
1299388/1500000: episode: 613, duration: 20.390s, episode steps: 3397, steps per second: 167, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1301293/1500000: episode: 614, duration: 11.520s, episode steps: 1905, steps per second: 165, episode reward: 17.000, mean reward: 0.009 [-1.000, 1.000], mean action:
1303280/1500000: episode: 615, duration: 12.005s, episode steps: 1987, steps per second: 166, episode reward: 17.000, mean reward: 0.009 [-1.000, 1.000], mean action:
1305528/1500000: episode: 616, duration: 13.447s, episode steps: 2248, steps per second: 167, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean action:
1307487/1500000: episode: 617, duration: 11.715s, episode steps: 1959, steps per second: 167, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean action:
1309669/1500000: episode: 618, duration: 13.151s, episode steps: 2182, steps per second: 166, episode reward: 18.000, mean reward: 0.008 [-1.000, 1.000], mean action:
1312098/1500000: episode: 619, duration: 14.609s, episode steps: 2429, steps per second: 166, episode reward: 11.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1315043/1500000: episode: 620, duration: 17.930s, episode steps: 2945, steps per second: 164, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1317435/1500000: episode: 621, duration: 14.473s, episode steps: 2392, steps per second: 165, episode reward: 11.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1319957/1500000: episode: 622, duration: 15.462s, episode steps: 2522, steps per second: 163, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1322078/1500000: episode: 623, duration: 12.853s, episode steps: 2121, steps per second: 165, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action:
1323781/1500000: episode: 624, duration: 10.284s, episode steps: 1703, steps per second: 166, episode reward: 21.000, mean reward: 0.012 [0.000, 1.000], mean action:
1325895/1500000: episode: 625, duration: 12.733s, episode steps: 2114, steps per second: 166, episode reward: -9.000, mean reward: -0.004 [-1.000, 1.000], mean action:
1327943/1500000: episode: 626, duration: 12.542s, episode steps: 2048, steps per second: 163, episode reward: 13.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1330628/1500000: episode: 627, duration: 16.292s, episode steps: 2685, steps per second: 165, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1332794/1500000: episode: 628, duration: 13.035s, episode steps: 2166, steps per second: 166, episode reward: 18.000, mean reward: 0.008 [-1.000, 1.000], mean action:
1335083/1500000: episode: 629, duration: 13.673s, episode steps: 2289, steps per second: 167, episode reward: 11.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1337783/1500000: episode: 630, duration: 16.206s, episode steps: 2700, steps per second: 167, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1339538/1500000: episode: 631, duration: 10.382s, episode steps: 1755, steps per second: 169, episode reward: 19.000, mean reward: 0.011 [-1.000, 1.000], mean action:
1341780/1500000: episode: 632, duration: 13.408s, episode steps: 2242, steps per second: 167, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1343699/1500000: episode: 633, duration: 11.576s, episode steps: 1919, steps per second: 166, episode reward: 17.000, mean reward: 0.009 [-1.000, 1.000], mean action:
1346330/1500000: episode: 634, duration: 15.732s, episode steps: 2631, steps per second: 167, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1348953/1500000: episode: 635, duration: 15.744s, episode steps: 2623, steps per second: 167, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1352098/1500000: episode: 636, duration: 18.803s, episode steps: 3145, steps per second: 167, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean action:
1354449/1500000: episode: 637, duration: 14.124s, episode steps: 2351, steps per second: 166, episode reward: 9.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1356927/1500000: episode: 638, duration: 14.741s, episode steps: 2478, steps per second: 168, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1359400/1500000: episode: 639, duration: 14.836s, episode steps: 2473, steps per second: 167, episode reward: 9.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1361311/1500000: episode: 640, duration: 11.438s, episode steps: 1911, steps per second: 167, episode reward: 19.000, mean reward: 0.010 [-1.000, 1.000], mean action:
1364329/1500000: episode: 641, duration: 18.174s, episode steps: 3018, steps per second: 166, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean action:
1366888/1500000: episode: 642, duration: 15.348s, episode steps: 2559, steps per second: 167, episode reward: 9.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1369259/1500000: episode: 643, duration: 14.382s, episode steps: 2371, steps per second: 165, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1370273/1500000: episode: 644, duration: 6.024s, episode steps: 1014, steps per second: 168, episode reward: -21.000, mean reward: -0.021 [-1.000, 0.000], mean action:
1372145/1500000: episode: 645, duration: 11.464s, episode steps: 1872, steps per second: 163, episode reward: -12.000, mean reward: -0.006 [-1.000, 1.000], mean action:
1374517/1500000: episode: 646, duration: 14.229s, episode steps: 2372, steps per second: 167, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1377000/1500000: episode: 647, duration: 14.786s, episode steps: 2483, steps per second: 168, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1379886/1500000: episode: 648, duration: 17.338s, episode steps: 2886, steps per second: 166, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1382116/1500000: episode: 649, duration: 13.384s, episode steps: 2230, steps per second: 167, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1384124/1500000: episode: 650, duration: 11.779s, episode steps: 2008, steps per second: 170, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean action:
1386582/1500000: episode: 651, duration: 14.631s, episode steps: 2458, steps per second: 168, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1388985/1500000: episode: 652, duration: 14.379s, episode steps: 2403, steps per second: 167, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1391301/1500000: episode: 653, duration: 13.903s, episode steps: 2316, steps per second: 167, episode reward: 13.000, mean reward: 0.006 [-1.000, 1.000], mean action:

```

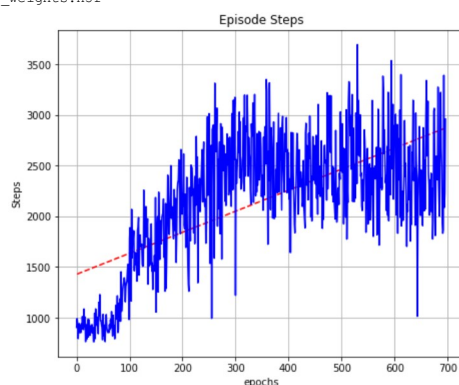
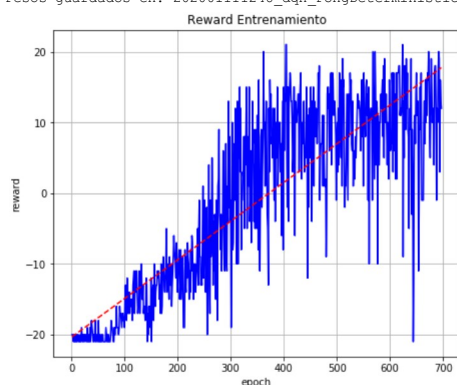
```

1393209/1500000: episode: 654, duration: 11.449s, episode steps: 1908, steps per second: 167, episode reward: -11.000, mean reward: -0.006 [-1.000, 1.000], mean acti
1396083/1500000: episode: 655, duration: 17.179s, episode steps: 2874, steps per second: 167, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean acti
1398868/1500000: episode: 656, duration: 16.521s, episode steps: 2785, steps per second: 169, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1400922/1500000: episode: 657, duration: 12.299s, episode steps: 2054, steps per second: 167, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean actor
1403476/1500000: episode: 658, duration: 15.213s, episode steps: 2554, steps per second: 168, episode reward: 9.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1405546/1500000: episode: 659, duration: 12.378s, episode steps: 2070, steps per second: 167, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean actor
1408493/1500000: episode: 660, duration: 17.606s, episode steps: 2947, steps per second: 167, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
1411833/1500000: episode: 661, duration: 20.071s, episode steps: 3340, steps per second: 166, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
1414428/1500000: episode: 662, duration: 15.518s, episode steps: 2595, steps per second: 167, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
1417148/1500000: episode: 663, duration: 16.413s, episode steps: 2720, steps per second: 166, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
1420283/1500000: episode: 664, duration: 18.782s, episode steps: 3135, steps per second: 167, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actor
1422704/1500000: episode: 665, duration: 14.862s, episode steps: 2421, steps per second: 163, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean actor
1425017/1500000: episode: 666, duration: 13.901s, episode steps: 2313, steps per second: 166, episode reward: 7.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1427029/1500000: episode: 667, duration: 11.894s, episode steps: 2012, steps per second: 169, episode reward: 16.000, mean reward: 0.008 [-1.000, 1.000], mean actor
1429389/1500000: episode: 668, duration: 14.221s, episode steps: 2360, steps per second: 166, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean actor
1431579/1500000: episode: 669, duration: 13.339s, episode steps: 2190, steps per second: 164, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean actor
1434004/1500000: episode: 670, duration: 15.156s, episode steps: 2425, steps per second: 160, episode reward: 11.000, mean reward: 0.005 [-1.000, 1.000], mean actor
1436047/1500000: episode: 671, duration: 12.563s, episode steps: 2043, steps per second: 163, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean actor
1438323/1500000: episode: 672, duration: 13.958s, episode steps: 2276, steps per second: 163, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean actor
1440099/1500000: episode: 673, duration: 10.853s, episode steps: 1776, steps per second: 164, episode reward: 20.000, mean reward: 0.011 [-1.000, 1.000], mean actor
1442079/1500000: episode: 674, duration: 12.407s, episode steps: 1980, steps per second: 160, episode reward: 17.000, mean reward: 0.009 [-1.000, 1.000], mean actor
1445116/1500000: episode: 675, duration: 18.741s, episode steps: 3037, steps per second: 162, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1448181/1500000: episode: 676, duration: 18.921s, episode steps: 3065, steps per second: 162, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1450684/1500000: episode: 677, duration: 15.406s, episode steps: 2503, steps per second: 162, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean actor
1452494/1500000: episode: 678, duration: 11.225s, episode steps: 1810, steps per second: 161, episode reward: 19.000, mean reward: 0.010 [-1.000, 1.000], mean actor
1454844/1500000: episode: 679, duration: 14.468s, episode steps: 2350, steps per second: 162, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean actor
1457438/1500000: episode: 680, duration: 16.038s, episode steps: 2594, steps per second: 162, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean actor
1459954/1500000: episode: 681, duration: 15.391s, episode steps: 2516, steps per second: 163, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean actor
1462315/1500000: episode: 682, duration: 14.364s, episode steps: 2361, steps per second: 164, episode reward: 11.000, mean reward: 0.005 [-1.000, 1.000], mean actor
1465589/1500000: episode: 683, duration: 19.669s, episode steps: 3274, steps per second: 166, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1467949/1500000: episode: 684, duration: 14.236s, episode steps: 2360, steps per second: 166, episode reward: 17.000, mean reward: 0.007 [-1.000, 1.000], mean actor
1470078/1500000: episode: 685, duration: 12.614s, episode steps: 2129, steps per second: 169, episode reward: 18.000, mean reward: 0.008 [-1.000, 1.000], mean actor
1472079/1500000: episode: 686, duration: 11.922s, episode steps: 2001, steps per second: 168, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean actor
1475090/1500000: episode: 687, duration: 18.281s, episode steps: 3011, steps per second: 165, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean actor
1478312/1500000: episode: 688, duration: 19.412s, episode steps: 3222, steps per second: 166, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
1480984/1500000: episode: 689, duration: 16.088s, episode steps: 2672, steps per second: 166, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean actor
1483479/1500000: episode: 690, duration: 15.121s, episode steps: 2495, steps per second: 165, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean actor
1485409/1500000: episode: 691, duration: 11.626s, episode steps: 1930, steps per second: 166, episode reward: 17.000, mean reward: 0.009 [-1.000, 1.000], mean actor
1487242/1500000: episode: 692, duration: 11.174s, episode steps: 1833, steps per second: 164, episode reward: 20.000, mean reward: 0.011 [-1.000, 1.000], mean actor
1489107/1500000: episode: 693, duration: 11.273s, episode steps: 1865, steps per second: 165, episode reward: 19.000, mean reward: 0.010 [-1.000, 1.000], mean actor
1492498/1500000: episode: 694, duration: 20.317s, episode steps: 3391, steps per second: 167, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1494585/1500000: episode: 695, duration: 12.632s, episode steps: 2087, steps per second: 165, episode reward: 16.000, mean reward: 0.008 [-1.000, 1.000], mean actor
1496819/1500000: episode: 696, duration: 13.373s, episode steps: 2234, steps per second: 167, episode reward: 13.000, mean reward: 0.006 [-1.000, 1.000], mean actor
1499778/1500000: episode: 697, duration: 17.796s, episode steps: 2959, steps per second: 166, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean actor

```

done, took 9042.429 seconds

Pesos guardados en: 202001111248_dqn_PongDeterministic-v0_weights.h5f



Testing for 10 episodes ...

```

Episode 1: reward: 20.000, steps: 1703
Episode 2: reward: 18.000, steps: 2125
Episode 3: reward: 14.000, steps: 2599
Episode 4: reward: 18.000, steps: 2191
Episode 5: reward: 20.000, steps: 1988
Episode 6: reward: 20.000, steps: 1862
Episode 7: reward: 11.000, steps: 2703
Episode 8: reward: 20.000, steps: 1818
Episode 9: reward: 17.000, steps: 2181
Episode 10: reward: 19.000, steps: 2008
mean reward: 17.7
mean_steps: 2117.8
(17.7, 2117.8)

```