

Training for 3500000 steps ...

8822/3500000:	episode: 1,	duration: 6.32s,	episode steps: 822,	steps per second: 130,	episode reward: -21.000,	mean reward: -0.026	[-1.000, 0.000],	mean action:
1646/3500000:	episode: 2,	duration: 4.462s,	episode steps: 824,	steps per second: 185,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
2526/3500000:	episode: 3,	duration: 4.744s,	episode steps: 880,	steps per second: 186,	episode reward: -21.000,	mean reward: -0.024	[-1.000, 0.000],	mean action:
3447/3500000:	episode: 4,	duration: 5.016s,	episode steps: 921,	steps per second: 184,	episode reward: -20.000,	mean reward: -0.022	[-1.000, 1.000],	mean action:
4271/3500000:	episode: 5,	duration: 4.642s,	episode steps: 824,	steps per second: 178,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
5035/3500000:	episode: 6,	duration: 4.150s,	episode steps: 764,	steps per second: 184,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
5859/3500000:	episode: 7,	duration: 4.443s,	episode steps: 824,	steps per second: 185,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
6870/3500000:	episode: 8,	duration: 5.516s,	episode steps: 1011,	steps per second: 183,	episode reward: -20.000,	mean reward: -0.020	[-1.000, 1.000],	mean action:
7670/3500000:	episode: 9,	duration: 4.479s,	episode steps: 800,	steps per second: 179,	episode reward: -21.000,	mean reward: -0.026	[-1.000, 0.000],	mean action:
8462/3500000:	episode: 10,	duration: 4.398s,	episode steps: 792,	steps per second: 180,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
9347/3500000:	episode: 11,	duration: 4.853s,	episode steps: 885,	steps per second: 182,	episode reward: -21.000,	mean reward: -0.024	[-1.000, 0.000],	mean action:
10305/3500000:	episode: 12,	duration: 4.572s,	episode steps: 958,	steps per second: 112,	episode reward: -20.000,	mean reward: -0.021	[-1.000, 1.000],	mean action:
11200/3500000:	episode: 13,	duration: 11.179s,	episode steps: 895,	steps per second: 80,	episode reward: -20.000,	mean reward: -0.022	[-1.000, 1.000],	mean action:
12237/3500000:	episode: 14,	duration: 13.175s,	episode steps: 1037,	steps per second: 79,	episode reward: -20.000,	mean reward: -0.019	[-1.000, 1.000],	mean action:
13149/3500000:	episode: 15,	duration: 11.582s,	episode steps: 912,	steps per second: 79,	episode reward: -21.000,	mean reward: -0.023	[-1.000, 0.000],	mean action:
13991/3500000:	episode: 16,	duration: 10.579s,	episode steps: 842,	steps per second: 80,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
14834/3500000:	episode: 17,	duration: 10.986s,	episode steps: 843,	steps per second: 77,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
15658/3500000:	episode: 18,	duration: 10.459s,	episode steps: 824,	steps per second: 79,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
16557/3500000:	episode: 19,	duration: 11.677s,	episode steps: 899,	steps per second: 77,	episode reward: -21.000,	mean reward: -0.023	[-1.000, 0.000],	mean action:
17559/3500000:	episode: 20,	duration: 12.937s,	episode steps: 1002,	steps per second: 77,	episode reward: -21.000,	mean reward: -0.021	[-1.000, 0.000],	mean action:
18506/3500000:	episode: 21,	duration: 12.091s,	episode steps: 947,	steps per second: 78,	episode reward: -20.000,	mean reward: -0.021	[-1.000, 1.000],	mean action:
19497/3500000:	episode: 22,	duration: 12.633s,	episode steps: 991,	steps per second: 78,	episode reward: -21.000,	mean reward: -0.021	[-1.000, 0.000],	mean action:
20289/3500000:	episode: 23,	duration: 10.104s,	episode steps: 792,	steps per second: 78,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
21251/3500000:	episode: 24,	duration: 12.241s,	episode steps: 962,	steps per second: 79,	episode reward: -21.000,	mean reward: -0.022	[-1.000, 0.000],	mean action:
22122/3500000:	episode: 25,	duration: 11.003s,	episode steps: 871,	steps per second: 79,	episode reward: -21.000,	mean reward: -0.024	[-1.000, 0.000],	mean action:
23009/3500000:	episode: 26,	duration: 11.172s,	episode steps: 887,	steps per second: 79,	episode reward: -20.000,	mean reward: -0.023	[-1.000, 1.000],	mean action:
23801/3500000:	episode: 27,	duration: 10.074s,	episode steps: 792,	steps per second: 79,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
24662/3500000:	episode: 28,	duration: 10.772s,	episode steps: 861,	steps per second: 80,	episode reward: -20.000,	mean reward: -0.023	[-1.000, 1.000],	mean action:
25642/3500000:	episode: 29,	duration: 12.143s,	episode steps: 980,	steps per second: 81,	episode reward: -20.000,	mean reward: -0.020	[-1.000, 1.000],	mean action:
26606/3500000:	episode: 30,	duration: 11.928s,	episode steps: 964,	steps per second: 81,	episode reward: -20.000,	mean reward: -0.021	[-1.000, 1.000],	mean action:
27446/3500000:	episode: 31,	duration: 10.524s,	episode steps: 840,	steps per second: 80,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
28322/3500000:	episode: 32,	duration: 11.218s,	episode steps: 876,	steps per second: 78,	episode reward: -21.000,	mean reward: -0.024	[-1.000, 0.000],	mean action:
29200/3500000:	episode: 33,	duration: 11.447s,	episode steps: 878,	steps per second: 77,	episode reward: -20.00			

10445/3500000:	episode: 109,	duration: 14.498s,	episode steps: 887,	steps per second: 77,	episode reward: -21.000,	mean reward: -0.024 [-1.000, 0.000],	mean action
101330/3500000:	episode: 110,	duration: 11.110s,	episode steps: 885,	steps per second: 80,	episode reward: -21.000,	mean reward: -0.024 [-1.000, 0.000],	mean action
102401/3500000:	episode: 111,	duration: 13.643s,	episode steps: 1071,	steps per second: 79,	episode reward: -20.000,	mean reward: -0.019 [-1.000, 1.000],	mean action
103782/3500000:	episode: 112,	duration: 17.393s,	episode steps: 1381,	steps per second: 79,	episode reward: -19.000,	mean reward: -0.014 [-1.000, 1.000],	mean action
105129/3500000:	episode: 113,	duration: 17.383s,	episode steps: 1347,	steps per second: 77,	episode reward: -20.000,	mean reward: -0.015 [-1.000, 1.000],	mean action
106488/3500000:	episode: 114,	duration: 17.592s,	episode steps: 1359,	steps per second: 77,	episode reward: -21.000,	mean reward: -0.015 [-1.000, 0.000],	mean action
107737/3500000:	episode: 115,	duration: 16.518s,	episode steps: 1249,	steps per second: 76,	episode reward: -20.000,	mean reward: -0.016 [-1.000, 1.000],	mean action
108656/3500000:	episode: 116,	duration: 12.096s,	episode steps: 919,	steps per second: 76,	episode reward: -20.000,	mean reward: -0.022 [-1.000, 1.000],	mean action
109991/3500000:	episode: 117,	duration: 17.357s,	episode steps: 1335,	steps per second: 77,	episode reward: -18.000,	mean reward: -0.013 [-1.000, 1.000],	mean action
111942/3500000:	episode: 118,	duration: 25.985s,	episode steps: 1951,	steps per second: 75,	episode reward: -17.000,	mean reward: -0.009 [-1.000, 1.000],	mean action
113300/3500000:	episode: 119,	duration: 17.349s,	episode steps: 1358,	steps per second: 78,	episode reward: -19.000,	mean reward: -0.014 [-1.000, 1.000],	mean action
114610/3500000:	episode: 120,	duration: 16.788s,	episode steps: 1310,	steps per second: 78,	episode reward: -19.000,	mean reward: -0.015 [-1.000, 1.000],	mean action
115960/3500000:	episode: 121,	duration: 17.132s,	episode steps: 1350,	steps per second: 79,	episode reward: -19.000,	mean reward: -0.014 [-1.000, 1.000],	mean action
117139/3500000:	episode: 122,	duration: 14.789s,	episode steps: 1179,	steps per second: 80,	episode reward: -21.000,	mean reward: -0.018 [-1.000, 0.000],	mean action
118494/3500000:	episode: 123,	duration: 17.212s,	episode steps: 1355,	steps per second: 79,	episode reward: -20.000,	mean reward: -0.015 [-1.000, 1.000],	mean action
119713/3500000:	episode: 124,	duration: 15.376s,	episode steps: 1219,	steps per second: 79,	episode reward: -20.000,	mean reward: -0.016 [-1.000, 1.000],	mean action
120627/3500000:	episode: 125,	duration: 11.554s,	episode steps: 914,	steps per second: 79,	episode reward: -21.000,	mean reward: -0.023 [-1.000, 0.000],	mean action
121623/3500000:	episode: 126,	duration: 12.498s,	episode steps: 996,	steps per second: 80,	episode reward: -21.000,	mean reward: -0.021 [-1.000, 0.000],	mean action
122601/3500000:	episode: 127,	duration: 17.277s,	episode steps: 1378,	steps per second: 80,	episode reward: -17.000,	mean reward: -0.012 [-1.000, 1.000],	mean action
124022/3500000:	episode: 128,	duration: 12.818s,	episode steps: 1021,	steps per second: 80,	episode reward: -20.000,	mean reward: -0.020 [-1.000, 1.000],	mean action
125009/3500000:	episode: 129,	duration: 12.370s,	episode steps: 987,	steps per second: 80,	episode reward: -20.000,	mean reward: -0.020 [-1.000, 1.000],	mean action
125960/3500000:	episode: 130,	duration: 11.885s,	episode steps: 951,	steps per second: 80,	episode reward: -20.000,	mean reward: -0.021 [-1.000, 1.000],	mean action
127255/3500000:	episode: 131,	duration: 16.290s,	episode steps: 1295,	steps per second: 79,	episode reward: -18.000,	mean reward: -0.014 [-1.000, 1.000],	mean action
128565/3500000:	episode: 132,	duration: 16.501s,	episode steps: 1310,	steps per second: 79,	episode reward: -19.000,	mean reward: -0.015 [-1.000, 1.000],	mean action
129880/3500000:	episode: 133,	duration: 16.575s,	episode steps: 1315,	steps per second: 79,	episode reward: -19.000,	mean reward: -0.014 [-1.000, 1.000],	mean action
131875/3500000:	episode: 134,	duration: 25.071s,	episode steps: 1995,	steps per second: 80,	episode reward: -20.000,	mean reward: -0.010 [-1.000, 1.000],	mean action
133977/3500000:	episode: 135,	duration: 26.536s,	episode steps: 2102,	steps per second: 79,	episode reward: -17.000,	mean reward: -0.008 [-1.000, 1.000],	mean action
135561/3500000:	episode: 136,	duration: 20.086s,	episode steps: 1584,	steps per second: 79,	episode reward: -20.000,	mean reward: -0.013 [-1.000, 1.000],	mean action
137430/3500000:	episode: 137,	duration: 23.380s,	episode steps: 1869,	steps per second: 80,	episode reward: -18.000,	mean reward: -0.010 [-1.000, 1.000],	mean action
138715/3500000:	episode: 138,	duration: 18.335s,	episode steps: 1445,	steps per second: 79,	episode reward: -16.000,	mean reward: -0.011 [-1.000, 1.000],	mean action
140109/3500000:	episode: 139,	duration: 15.528s,	episode steps: 1234,	steps per second: 79,	episode reward: -20.000,	mean reward: -0.016 [-1.000, 1.000],	mean action
141585/3500000:	episode: 140,	duration: 18.562s,	episode steps: 1476,	steps per second: 80,	episode reward: -19.000,	mean reward: -0.013 [-1.000, 1.000],	mean action
142489/3500000:	episode: 141,	duration: 11.337s,	episode steps: 904,	steps per second: 80,	episode reward: -20.000,	mean reward: -0.022 [-1.000, 1.000],	mean action
143991/							

305557/3500000:	episode: 218,	duration: 22.084s,	episode steps: 1839,	steps per second: 83,	episode reward: -20.000,	mean reward: -0.011	[-1.000, 1.000],	mean actio
307846/3500000:	episode: 219,	duration: 27.223s,	episode steps: 2289,	steps per second: 84,	episode reward: -10.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
310284/3500000:	episode: 220,	duration: 28.325s,	episode steps: 2438,	steps per second: 86,	episode reward: -16.000,	mean reward: -0.007	[-1.000, 1.000],	mean actio
311854/3500000:	episode: 221,	duration: 18.278s,	episode steps: 1570,	steps per second: 86,	episode reward: -19.000,	mean reward: -0.012	[-1.000, 1.000],	mean actio
314780/3500000:	episode: 222,	duration: 35.292s,	episode steps: 2926,	steps per second: 83,	episode reward: -13.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
317339/3500000:	episode: 223,	duration: 29.521s,	episode steps: 2559,	steps per second: 87,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean actio
319926/3500000:	episode: 224,	duration: 32.719s,	episode steps: 2587,	steps per second: 79,	episode reward: -17.000,	mean reward: -0.007	[-1.000, 1.000],	mean actio
322324/3500000:	episode: 225,	duration: 31.100s,	episode steps: 2398,	steps per second: 77,	episode reward: -12.000,	mean reward: -0.005	[-1.000, 1.000],	mean actio
326021/3500000:	episode: 226,	duration: 47.200s,	episode steps: 3697,	steps per second: 78,	episode reward: -10.000,	mean reward: -0.003	[-1.000, 1.000],	mean actio
328646/3500000:	episode: 227,	duration: 32.981s,	episode steps: 2625,	steps per second: 80,	episode reward: -14.000,	mean reward: -0.005	[-1.000, 1.000],	mean actio
331018/3500000:	episode: 228,	duration: 30.267s,	episode steps: 2372,	steps per second: 78,	episode reward: -16.000,	mean reward: -0.007	[-1.000, 1.000],	mean actio
334216/3500000:	episode: 229,	duration: 39.989s,	episode steps: 3198,	steps per second: 80,	episode reward: -12.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
336854/3500000:	episode: 230,	duration: 33.487s,	episode steps: 2638,	steps per second: 79,	episode reward: -12.000,	mean reward: -0.005	[-1.000, 1.000],	mean actio
339531/3500000:	episode: 231,	duration: 34.079s,	episode steps: 2677,	steps per second: 79,	episode reward: -11.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
342997/3500000:	episode: 232,	duration: 44.334s,	episode steps: 3466,	steps per second: 78,	episode reward: -6.000,	mean reward: -0.002	[-1.000, 1.000],	mean actio
345801/3500000:	episode: 233,	duration: 35.585s,	episode steps: 2804,	steps per second: 79,	episode reward: -12.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
347395/3500000:	episode: 234,	duration: 20.179s,	episode steps: 1594,	steps per second: 79,	episode reward: -18.000,	mean reward: -0.011	[-1.000, 1.000],	mean actio
350827/3500000:	episode: 235,	duration: 44.021s,	episode steps: 3432,	steps per second: 78,	episode reward: -6.000,	mean reward: -0.002	[-1.000, 1.000],	mean actio
352897/3500000:	episode: 236,	duration: 26.428s,	episode steps: 2070,	steps per second: 78,	episode reward: -16.000,	mean reward: -0.008	[-1.000, 1.000],	mean actio
355992/3500000:	episode: 237,	duration: 39.625s,	episode steps: 3095,	steps per second: 78,	episode reward: -7.000,	mean reward: -0.002	[-1.000, 1.000],	mean actio
358850/3500000:	episode: 238,	duration: 36.961s,	episode steps: 2858,	steps per second: 77,	episode reward: -16.000,	mean reward: -0.006	[-1.000, 1.000],	mean actio
360818/3500000:	episode: 239,	duration: 24.894s,	episode steps: 1968,	steps per second: 79,	episode reward: -19.000,	mean reward: -0.010	[-1.000, 1.000],	mean actio
362783/3500000:	episode: 240,	duration: 24.746s,	episode steps: 1965,	steps per second: 79,	episode reward: -18.000,	mean reward: -0.009	[-1.000, 1.000],	mean actio
365299/3500000:	episode: 241,	duration: 32.510s,	episode steps: 2516,	steps per second: 77,	episode reward: -14.000,	mean reward: -0.006	[-1.000, 1.000],	mean actio
367573/3500000:	episode: 242,	duration: 28.633s,	episode steps: 2274,	steps per second: 79,	episode reward: -18.000,	mean reward: -0.008	[-1.000, 1.000],	mean actio
370733/3500000:	episode: 243,	duration: 40.353s,	episode steps: 3160,	steps per second: 78,	episode reward: -5.000,	mean reward: -0.002	[-1.000, 1.000],	mean actio
372543/3500000:	episode: 244,	duration: 22.830s,	episode steps: 1810,	steps per second: 79,	episode reward: -18.000,	mean reward: -0.010	[-1.000, 1.000],	mean actio
375213/3500000:	episode: 245,	duration: 34.090s,	episode steps: 2670,	steps per second: 78,	episode reward: -11.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
377664/3500000:	episode: 246,	duration: 31.037s,	episode steps: 2451,	steps per second: 79,	episode reward: -10.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
380056/3500000:	episode: 247,	duration: 29.854s,	episode steps: 2392,	steps per second: 80,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean actio
382471/3500000:	episode: 248,	duration: 28.913s,	episode steps: 2415,	steps per second: 84,	episode reward: -15.000,	mean reward: -0.006	[-1.000, 1.000],	mean actio
384450/3500000:	episode: 249,	duration: 23.274s,	episode steps: 1979,	steps per second: 85,	episode reward: -16.000,	mean reward: -0.008	[-1.000, 1.000],	mean actio
386280/3500000:	episode: 250,	duration: 21.251s,	episode steps: 1830,	steps per second: 86,	episode reward: -18.000,	mean reward: -0.010	[-1.000, 1.000],	mean actio
388484/3500000:	episode: 251,	duration: 25.490s,	episode steps: 2204,	steps per second: 86,	episode reward: -15.000,	mean reward: -0.007	[-1.000, 1.000],	mean actio
391073/3500000:	episode: 252,	duration: 29.890s,	episode steps: 2589,	steps per second: 87,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean actio
394167/3500000:	episode: 253,	duration: 35.569s,	episode steps: 3094,	steps per second: 87,	episode reward: -9.000,	mean reward: -0.003	[-1.000, 1.000],	mean actio
396764/3500000:	episode: 254,	duration: 30.020s,	episode steps: 2597,	steps per second: 87,	episode reward: -16.000,	mean reward: -0.006	[-1.000, 1.000],	mean actio
398992/3500000:	episode: 255,	duration: 25.621s,	episode steps: 2228,	steps per second: 87,	episode reward: -14.000,	mean reward: -0.006	[-1.000, 1.000],	mean actio
401834/3500000:	episode: 256,	duration: 32.822s,	episode steps: 2842,	steps per second: 87,	episode reward: -12.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
404026/3500000:	episode: 257,	duration: 25.088s,	episode steps: 2192,	steps per second: 87,	episode reward: -15.000,	mean reward: -0.007	[-1.000, 1.000],	mean actio
406732/3500000:	episode: 258,	duration: 32.543s,	episode steps: 2706,	steps per second: 83,	episode reward: -11.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
409538/3500000:	episode: 259,	duration: 33.546s,	episode steps: 2806,	steps per second: 84,	episode reward: -15.000,	mean reward: -0.005	[-1.000, 1.000],	mean actio
413140/3500000:	episode: 260,	duration: 41.525s,	episode steps: 3602,	steps per second: 87,	episode reward: -8.000,	mean reward: -0.002	[-1.000, 1.000],	mean actio
415095/3500000:	episode: 261,	duration: 22.423s,	episode steps: 1955,	steps per second: 87,	episode reward: -17.000,	mean reward: -0.009	[-1.000, 1.000],	mean actio
417941/3500000:	episode: 262,	duration: 32.865s,	episode steps: 2846,	steps per second: 87,	episode reward: -10.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
420467/3500000:	episode: 263,	duration: 29.021s,	episode steps: 2526,	steps per second: 87,	episode reward: -11.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
422969/3500000:	episode: 264,	duration: 28.815s,	episode steps: 2502,	steps per second: 87,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean actio
425138/3500000:	episode: 265,	duration: 25.011s,	episode steps: 2169,	steps per second: 87,	episode reward: -11.000,	mean reward: -0.005	[-1.000, 1.000],	mean actio
428379/3500000:	episode: 266,	duration: 37.334s,	episode steps: 3241,	steps per second: 87,	episode reward: -6.000,	mean reward: -0.002	[-1.000, 1.000],	mean actio
431778/3500000:	episode: 267,	duration: 39.962s,	episode steps: 3399,	steps per second: 85,	episode reward: -9.000,	mean reward: -0.003	[-1.000, 1.000],	mean actio
435060/3500000:	episode: 268,	duration: 39.350s,	episode steps: 3282,	steps per second: 83,	episode reward: -13.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
437402/3500000:	episode: 269,	duration: 27.197s,	episode steps: 2342,	steps per second: 86,	episode reward: -14.000,	mean reward: -0.006	[-1.000, 1.000],	mean actio
440101/3500000:	episode: 270,	duration: 31.043s,	episode steps: 2699,	steps per second: 87,	episode reward: -12.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
442510/3500000:	episode: 271,	duration: 27.788s,	episode steps: 2409,	steps per second: 87,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean actio
445045/3500000:	episode: 272,	duration: 29.170s,	episode steps: 2535,	steps per second: 87,	episode reward: -14.000,	mean reward: -0.006	[-1.000, 1.000],	mean actio
448089/3500000:	episode: 273,	duration: 34.947s,	episode steps: 3044,	steps per second: 87,	episode reward: -9.000,	mean reward: -0.003	[-1.000, 1.000],	mean actio
451376/3500000:	episode: 274,	duration: 37.908s,	episode steps: 3287,	steps per second: 87,	episode reward: -9.000,	mean reward: -0.003	[-1.000, 1.000],	mean actio
454436/3500000:	episode: 275,	duration: 35.163s,	episode steps: 3060,	steps per second: 87,	episode reward: -6.000,	mean reward: -0.002	[-1.000, 1.000],	mean actio
457910/3500000:	episode: 276,	duration: 39.945s,	episode steps: 3474,	steps per second: 87,	episode reward: -6.000,	mean reward: -0.002	[-1.000, 1.000],	mean actio
461278/3500000:	episode: 277,	duration: 38.638s,	episode steps: 3368,	steps per second: 87,	episode reward: -9.000,	mean reward: -0.003	[-1.000, 1.000],	mean actio
463466/3500000:	episode: 278,	duration: 25.184s,	episode steps: 2188,	steps per second: 87,	episode reward: -14.000,	mean reward: -0.006	[-1.000, 1.000],	mean actio
465669/3500000:	episode: 279,	duration: 25.492s,	episode steps: 2203,	steps per second: 86,	episode reward: -15.000,	mean reward: -0.007	[-1.000, 1.000],	mean actio
468488/3500000:	episode: 280,	duration: 32.374s,	episode steps: 2819,	steps per second: 87,	episode reward: -11.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
472459/3500000:	episode: 281,	duration: 45.791s,	episode steps: 3971,	steps per second: 87,	episode reward: -6.000,	mean reward: -0.002	[-1.000, 1.000],	mean actio
475025/3500000:	episode: 282,	duration: 29.748s,	episode steps: 2566,	steps per second: 86,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean actio
477826/3500000:	episode: 283,	duration: 32.104s,	episode steps: 2801,	steps per second: 87,	episode reward: -10.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
480399/3500000:	episode: 284,	duration: 29.650s,	episode steps: 2573,	steps per second: 87,	episode reward: -12.000,	mean reward: -0.005	[-1.000, 1.000],	mean actio
483076/3500000:	episode: 285,	duration: 30.821s,	episode steps: 2677,	steps per second: 87,	episode reward: -14.000,	mean reward: -0.005	[-1.000, 1.000],	mean actio
486823/3500000:	episode: 286,	duration: 42.937s,	episode steps: 3747,	steps per second: 87,	episode reward: -6.000,	mean reward: -0.002	[-1.000, 1.000],	mean actio
489236/3500000:	episode: 287,	duration: 27.786s,	episode steps: 2413,	steps per second: 87,	episode reward: -15.000,	mean reward: -0.006	[-1.000, 1.000],	mean actio
492345/3500000:	episode: 288,	duration: 35.731s,	episode steps: 3109,	steps per second: 87,	episode reward: -9.000,	mean reward: -0.003	[-1.000, 1.000],	mean actio
495949/3500000:	episode: 289,	duration: 41.460s,	episode steps: 3604,	steps per second: 87,	episode reward: -4.000,	mean reward: -0.001	[-1.000, 1.000],	mean actio
498572/3500000:	episode: 290,	duration: 30.039s,	episode steps: 2623,	steps per second: 87,	episode reward: -11.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
501854/3500000:	episode: 291,	duration: 41.483s,	episode steps: 3282,	steps per second: 79,	episode reward: -5.000,	mean reward: -0.002	[-1.000, 1.000],	mean actio
504026/3500000:	episode: 292,	duration: 27.490s,	episode steps: 2172,	steps per second: 79,	episode reward: -13.000,	mean reward: -0.006	[-1.000, 1.000],	mean actio
506081/3500000:	episode: 293,	duration: 26.086s,	episode steps: 2055,	steps per second: 79,	episode reward: -13.000,	mean reward: -0.006	[-1.000, 1.000],	mean actio
508843/3500000:	episode: 294,	duration: 34.961s,	episode steps: 2762,	steps per second: 79,	episode reward: -12.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
511699/3500000:	episode: 295,	duration: 35.986s,	episode steps: 2856,	steps per second: 79,	episode reward: -12.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
515031/3500000:	episode: 296,	duration: 42.153s,	episode steps: 3332,	steps per second: 79,	episode reward: -9.000,	mean reward: -0.003	[-1.000, 1.000],	mean actio
517605/3500000:	episode: 297,	duration: 32.946s,	episode steps: 2574,	steps per second: 78,	episode reward: -14.000,	mean reward: -0.005	[-1.000, 1.000],	mean actio
520190/3500000:	episode: 298,	duration: 33.044s,	episode steps: 2585,	steps per second: 78,	episode reward: -15.000,	mean reward: -0.006	[-1.000, 1.000],	mean actio
523358/3500000:	episode: 299,	duration: 40.253s,	episode steps: 3168,	steps per second: 79,	episode reward: -9.000,	mean reward: -0.003	[-1.000, 1.000],	mean actio
525894/3500000:	episode: 300,	duration: 31.811s,	episode steps: 2536,	steps per second: 80,	episode reward: -15.000,	mean reward: -0.006	[-1.000, 1.000],	mean actio
529130/3500000:	episode: 301,	duration: 41.558s,	episode steps: 3236,	steps per second: 78,	episode reward: -9.000,	mean reward: -0.003	[-1.000, 1.000],	mean actio
532131/3500000:	episode: 302,	duration: 38.042s,	episode steps: 3001,	steps per second: 79,	episode reward: -12.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
536288/3500000:	episode: 303,	duration: 52.305s,	episode steps: 4157,	steps per second: 79,	episode reward: -3.000,	mean reward: -0.001	[-1.000, 1.000],	mean actio
539502/3500000:	episode: 304,	duration: 41.511s,	episode steps: 3214,	steps per second: 77,	episode reward: -9.000,	mean reward: -0.003	[-1.000, 1.000],	mean actio
542157/3500000:	episode: 305,	duration: 33.723s,	episode steps: 2655,	steps per second: 79,	episode reward: -12.000,	mean reward: -0.005	[-1.000, 1.000],	mean actio
545019/3500000:	episode: 306,	duration: 36.316s,	episode steps: 2862,	steps per second: 79,	episode reward: -12.000,	mean reward: -0.004	[-1.000, 1.000],	mean actio
547243/3500000:	episode: 307,	duration: 28.340s,	episode steps: 2224,	steps per second: 78,	episode reward: -15.000,	mean reward: -0.007	[-1.000, 1.000],	mean actio
549775/3500000:	episode: 308,	duration: 32.392s,	episode steps: 2532,	steps per second: 78,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean actio
553748/3500000:	episode: 309,	duration: 51.078s,	episode steps: 3973,	steps per second: 78,	episode reward: -3.000,	mean reward: -0.001	[-1.000, 1.000],	mean actio
557210/3500000:	episode: 310,	duration: 43.884s,	episode steps: 3462,	steps per second: 79,	episode reward: -9.000,	mean reward: -0.003	[-1.000, 1.000],	mean actio
560710/3500000:	episode: 311,	duration: 44.325s,	episode steps: 3500,	steps per second: 79,	episode reward: -5.000,	mean reward: -0.001	[-1.000, 1.000],	mean actio

613902/3500000:	episode: 327,	duration: 32.015s,	episode steps: 2776,	steps per second: 87,	episode reward: -12.000,	mean reward: -0.004 [-1.000, 1.000],	mean action:
617207/3500000:	episode: 328,	duration: 38.053s,	episode steps: 3305,	steps per second: 87,	episode reward: -9.000,	mean reward: -0.003 [-1.000, 1.000],	mean action:
621195/3500000:	episode: 329,	duration: 46.665s,	episode steps: 3988,	steps per second: 85,	episode reward: 4.000,	mean reward: 0.001 [-1.000, 1.000],	mean action:
624980/3500000:	episode: 330,	duration: 45.506s,	episode steps: 3785,	steps per second: 83,	episode reward: 8.000,	mean reward: 0.002 [-1.000, 1.000],	mean action:
628496/3500000:	episode: 331,	duration: 40.435s,	episode steps: 3516,	steps per second: 87,	episode reward: -7.000,	mean reward: -0.002 [-1.000, 1.000],	mean action:
632273/3500000:	episode: 332,	duration: 43.521s,	episode steps: 3777,	steps per second: 87,	episode reward: -6.000,	mean reward: -0.002 [-1.000, 1.000],	mean action:
635576/3500000:	episode: 333,	duration: 37.982s,	episode steps: 3303,	steps per second: 87,	episode reward: -11.000,	mean reward: -0.003 [-1.000, 1.000],	mean action:
638837/3500000:	episode: 334,	duration: 37.607s,	episode steps: 3261,	steps per second: 87,	episode reward: -8.000,	mean reward: -0.002 [-1.000, 1.000],	mean action:
641571/3500000:	episode: 335,	duration: 31.429s,	episode steps: 2734,	steps per second: 87,	episode reward: -9.000,	mean reward: -0.003 [-1.000, 1.000],	mean action:
644651/3500000:	episode: 336,	duration: 35.367s,	episode steps: 3080,	steps per second: 87,	episode reward: -9.000,	mean reward: -0.003 [-1.000, 1.000],	mean action:
647359/3500000:	episode: 337,	duration: 32.169s,	episode steps: 2708,	steps per second: 84,	episode reward: -9.000,	mean reward: -0.003 [-1.000, 1.000],	mean action:
650510/3500000:	episode: 338,	duration: 39.497s,	episode steps: 3151,	steps per second: 80,	episode reward: -10.000,	mean reward: -0.003 [-1.000, 1.000],	mean action:
654746/3500000:	episode: 339,	duration: 54.438s,	episode steps: 4236,	steps per second: 78,	episode reward: -5.000,	mean reward: -0.001 [-1.000, 1.000],	mean action:
658165/3500000:	episode: 340,	duration: 43.109s,	episode steps: 3419,	steps per second: 79,	episode reward: -7.000,	mean reward: -0.002 [-1.000, 1.000],	mean action:
661935/3500000:	episode: 341,	duration: 47.941s,	episode steps: 3770,	steps per second: 79,	episode reward: -7.000,	mean reward: -0.002 [-1.000, 1.000],	mean action:
665783/3500000:	episode: 342,	duration: 48.593s,	episode steps: 3848,	steps per second: 79,	episode reward: 2.000,	mean reward: 0.001 [-1.000, 1.000],	mean action:
669390/3500000:	episode: 343,	duration: 45.665s,	episode steps: 3607,	steps per second: 79,	episode reward: -8.000,	mean reward: -0.002 [-1.000, 1.000],	mean action:
672626/3500000:	episode: 344,	duration: 40.804s,	episode steps: 3236,	steps per second: 79,	episode reward: -14.000,	mean reward: -0.004 [-1.000, 1.000],	mean action:
676098/3500000:	episode: 345,	duration: 43.513s,	episode steps: 3472,	steps per second: 80,	episode reward: -9.000,	mean reward: -0.003 [-1.000, 1.000],	mean action:
679173/3500000:	episode: 346,	duration: 38.652s,	episode steps: 3075,	steps per second: 80,	episode reward: -12.000,	mean reward: -0.004 [-1.000, 1.000],	mean action:
682294/3500000:	episode: 347,	duration: 39.264s,	episode steps: 3121,	steps per second: 79,	episode reward: -7.000,	mean reward: -0.002 [-1.000, 1.000],	mean action:
686225/3500000:	episode: 348,	duration: 49.302s,	episode steps: 3931,	steps per second: 80,	episode reward: -1.000,	mean reward: -0.000 [-1.000, 1.000],	mean action:
687742/3500000:	episode: 349,	duration: 18.959s,	episode steps: 1517,	steps per second: 80,	episode reward: -15.000,	mean reward: -0.010 [-1.000, 1.000],	mean action:
688924/3500000:	episode: 350,	duration: 14.847s,	episode steps: 1182,	steps per second: 80,	episode reward: -18.000,	mean reward: -0.015 [-1.000, 1.000],	mean action:
691015/3500000:	episode: 351,	duration: 26.016s,	episode steps: 2091,	steps per second: 80,	episode reward: -13.000,	mean reward: -0.006 [-1.000, 1.000],	mean action:
693771/3500000:	episode: 352,	duration: 34.437s,	episode steps: 2756,	steps per second: 80,	episode reward: -15.000,	mean reward: -0.005 [-1.000, 1.000],	mean action:
697279/3500000:	episode: 353,	duration: 43.830s,	episode steps: 3508,	steps per second: 80,	episode reward: -7.000,	mean reward: -0.002 [-1.000, 1.000],	mean action:
699648/3500000:	episode: 354,	duration: 29.502s,	episode steps: 2369,	steps per second: 80,	episode reward: -16.000,	mean reward: -0.007 [-1.000, 1.000],	mean action:
702556/3500000:	episode: 355,	duration: 36.303s,	episode steps: 2908,	steps per second: 80,	episode reward: 9.000,	mean reward: 0.003 [-1.000, 1.000],	mean action:
707447/3500000:	episode: 356,	duration: 60.888s,	episode steps: 4891,	steps per second: 80,	episode reward: -5.000,	mean reward: -0.001 [-1.000, 1.000],	mean action:
710894/3500000:	episode: 357,	duration: 42.858s,	episode steps: 3447,	steps per second: 80,	episode reward: -2.000,	mean reward: -0.001 [-1.000, 1.000],	mean action:
714556/3500000:	episode: 358,	duration: 45.746s,	episode steps: 3662,	steps per second: 80,	episode reward: -8.000,	mean reward: -0.002 [-1.000, 1.000],	mean action:
718044/3500000:	episode: 359,	duration: 43.454s,	episode steps: 3488,	steps per second: 80,	episode reward: -4.000,	mean reward: -0.001 [-1.000, 1.000],	mean action:
721119/3500000:	episode: 360,	duration: 38.422s,	episode steps: 3075,	steps per second: 80,	episode reward: -9.000,	mean reward: -0.003 [-1.000, 1.000],	mean action:
725074/3500000:	episode: 361,	duration: 49.585s,	episode steps: 3955,	steps per second: 80,	episode reward: -5.000,	mean reward: -0.001 [-1.000, 1.000],	mean action:
729157/3500000:	episode: 362,	duration: 50.832s,	episode steps: 4083,	steps per second: 80,	episode reward: -4.000,	mean reward: -0.001 [-1.000, 1.000],	mean action:
734242/3500000:	episode: 363,	duration: 63.629s,	episode steps: 5085,	steps per second: 80,	episode reward: 1.000,	mean reward: 0.000 [-1.000, 1.000],	mean action:
737484/3500000:	episode: 364,	duration: 40.362s,	episode steps: 3242,	steps per second: 80,	episode reward: -6.000,	mean reward: -0.002 [-1.000, 1.000],	mean action:
741478/3500000:	episode: 365,	duration: 49.554s,	episode steps: 3994,	steps per second: 81,	episode reward: -5.000,	mean reward: -0.001 [-1.000, 1.000],	mean action:
744976/3500000:	episode: 366,	duration: 43.476s,	episode steps: 3498,	steps per second: 80,	episode reward: 8.000,	mean reward: 0.002 [-1.000, 1.000],	mean action:
749017/3500000:	episode: 367,	duration: 50.419s,	episode steps: 4041,	steps per second: 80,	episode reward: -5.000,	mean reward: -0.001 [-1.000, 1.000],	mean action:
753563/3500000:	episode: 368,	duration: 56.578s,	episode steps: 4546,	steps per second: 80,	episode reward: 3.000,	mean reward: 0.001 [-1.000, 1.000],	mean action:
756734/3500000:	episode: 369,	duration: 39.498s,	episode steps: 3171,	steps per second: 80,	episode reward: 6.000,	mean reward: 0.002 [-1.000, 1.000],	mean action:
759411/3500000:	episode: 370,	duration: 33.348s,	episode steps: 2677,	steps per second: 80,	episode reward: -9.000,	mean reward: -0.003 [-1.000, 1.000],	mean action:
761872/3500000:	episode: 371,	duration: 30.865s,	episode steps: 2461,	steps per second: 80,	episode reward: -13.000,	mean reward: -0.005 [-1.000, 1.000],	mean action:
765318/3500000:	episode: 372,	duration: 42.955s,	episode steps: 3446,	steps per second: 80,	episode reward: -5.000,	mean reward: -0.001 [-1.000, 1.000],	mean action:
768648/3500000:	episode: 373,	duration: 41.439s,	episode steps: 3330,	steps per second: 80,	episode reward: 4.000,	mean reward: 0.001 [-1.000, 1.000],	mean action:
772066/3500000:	episode: 374,	duration: 42.684s,	episode steps: 3418,	steps per second: 80,	episode reward: 6.000,	mean reward: 0.002 [-1.000, 1.000],	mean action:
775786/3500000:	episode: 375,	duration: 46.076s,	episode steps: 3720,	steps per second: 81,	episode reward: 4.000,	mean reward: 0.001 [-1.000, 1.000],	mean action:
778799/3500000:	episode: 376,	duration: 37.262s,	episode steps: 3013,	steps per second: 81,	episode reward: 10.000,	mean reward: 0.003 [-1.000, 1.000],	mean action:
782886/3500000:	episode: 377,	duration: 51.000s,	episode steps: 4087,	steps per second: 80,	episode reward: 6.000,	mean reward: 0.001 [-1.000, 1.000],	mean action:
786957/3500000:	episode: 378,	duration: 50.112s,	episode steps: 4071,	steps per second: 81,	episode reward: -1.000,	mean reward: -0.000 [-1.000, 1.000],	mean action:
790680/3500000:	episode: 379,	duration: 46.139s,	episode steps: 3723,	steps per second: 81,	episode reward: -5.000,	mean reward: -0.001 [-1.000, 1.000],	mean action:
794419/3500000:	episode: 380,	duration: 46.044s,	episode steps: 3739,	steps per second: 81,	episode reward: -2.000,	mean reward: -0.001 [-1.000, 1.000],	mean action:
798440/3500000:	episode: 381,	duration: 49.679s,	episode steps: 4021,	steps per second: 81,	episode reward: -2.000,	mean reward: -0.000 [-1.000, 1.000],	mean action:
802661/3500000:	episode: 382,	duration: 52.052s,	episode steps: 4221,	steps per second: 81,	episode reward: -1.000,	mean reward: -0.000 [-1.000, 1.000],	mean action:
806161/3500000:	episode: 383,	duration: 43.183s,	episode steps: 3500,	steps per second: 81,	episode reward: -7.000,	mean reward: -0.002 [-1.000, 1.000],	mean action:
809623/3500000:	episode: 384,	duration: 42.826s,	episode steps: 3462,	steps per second: 81,	episode reward: 7.000,	mean reward: 0.002 [-1.000, 1.000],	mean action:
812848/3500000:	episode: 385,	duration: 39.831s,	episode steps: 3225,	steps per second: 81,	episode reward: -9.000,	mean reward: -0.003 [-1.000, 1.000],	mean action:
815541/3500000:	episode: 386,	duration: 33.168s,	episode steps: 2693,	steps per second: 81,	episode reward: -9.000,	mean reward: -0.003 [-1.000, 1.000],	mean action:
817805/3500000:	episode: 387,	duration: 27.977s,	episode steps: 2264,	steps per second: 81,	episode reward: -13.000,	mean reward: -0.006 [-1.000, 1.000],	mean action:
821430/3500000:	episode: 388,	duration: 44.610s,	episode steps: 3625,	steps per second: 81,	episode reward: -1.000,	mean reward: -0.000 [-1.000, 1.000],	mean action:
825265/3500000:	episode: 389,	duration: 47.145s,	episode steps: 3835,	steps per second: 81,	episode reward: -3.000,	mean reward: -0.001 [-1.000, 1.000],	mean action:
828759/3500000:	episode: 390,	duration: 43.696s,	episode steps: 3494,	steps per second: 80,	episode reward: -10.000,	mean reward: -0.003 [-1.000, 1.000],	mean action:
832650/3500000:	episode: 391,	duration: 48.217s,	episode steps: 3891,	steps per second: 81,	episode reward: 1.000,	mean reward: 0.000 [-1.000, 1.000],	mean action:
835641/3500000:	episode: 392,	duration: 37.950s,	episode steps: 2991,	steps per second: 79,	episode reward: 6.000,	mean reward: 0.002 [-1.000, 1.000],	mean action:
839784/3500000:	episode: 393,	duration: 52.389s,	episode steps: 4143,	steps per second: 79,	episode reward: 5.000,	mean reward: 0.001 [-1.000, 1.000],	mean action:
842750/3500000:	episode: 394,	duration: 36.819s,	episode steps: 2966,	steps per second: 81,	episode reward: 8.000,	mean reward: 0.003 [-1.000, 1.000],	mean action:
846486/3500000:	episode: 395,	duration: 46.178s,	episode steps: 3736,	steps per second: 81,	episode reward: 7.000,	mean reward: 0.002 [-1.000, 1.000],	mean action:
849702/3500000:	episode: 396,	duration: 39.879s,	episode steps: 3216,	steps per second: 81,	episode reward: 9.000,	mean reward: 0.003 [-1.000, 1.000],	mean action:
852187/3500000:	episode: 397,	duration: 31.635s,	episode steps: 2485,	steps per second: 79,	episode reward: -17.000,	mean reward: -0.007 [-1.000, 1.000],	mean action:
856329/3500000:	episode: 398,	duration: 53.444s,	episode steps: 4142,	steps per second: 78,	episode reward: 4.000,	mean reward: 0.001 [-1.000, 1.000],	mean action:
860211/3500000:	episode: 399,	duration: 49.542s,	episode steps: 3882,	steps per second: 78,	episode reward: -5.000,	mean reward: -0.001 [-1.000, 1.000],	mean action:
864353/3500000:	episode: 400,	duration: 52.747s,	episode steps: 4142,	steps per second: 79,	episode reward: 3.000,	mean reward: 0.001 [-1.000, 1.000],	mean action:
868324/3500000:	episode: 401,	duration: 50.609s,	episode steps: 3971,	steps per second: 78,	episode reward: 3.000,	mean reward: 0.001 [-1.000, 1.000],	mean action:
871811/3500000:	episode: 402,	duration: 43.786s,	episode steps: 3487,	steps per second: 80,	episode reward: 9.000,	mean reward: 0.003 [-1.000, 1.000],	mean action:
874758/3500000:	episode: 403,	duration: 36.887s,	episode steps: 2947,	steps per second: 80,	episode reward: 8.000,	mean reward: 0.003 [-1.000, 1.000],	mean action:
878667/3500000:	episode: 404,	duration: 50.061s,	episode steps: 3909,	steps per second: 78,	episode reward: 5.000,	mean reward: 0.001 [-1.000, 1.000],	mean action:
882849/3500000:	episode: 405,	duration: 52.540s,	episode steps: 4182,	steps per second: 80,	episode reward: 1.000,	mean reward: 0.000 [-1.000, 1.000],	mean action:
886612/3500000:	episode: 406,	duration: 47.149s,	episode steps: 3763,	steps per second: 80,	episode reward: 3.000,	mean reward: 0.001 [-1.000, 1.000],	mean action:
889873/3500000:	episode: 407,	duration: 41.347s,	episode steps: 3261,	steps per second: 79,	episode reward: -10.000,	mean reward: -0.003 [-1.000, 1.000],	mean action:
892572/3500000:	episode: 408,	duration: 34.376s,	episode steps: 2699,	steps per second: 79,	episode reward: 12.000,	mean reward: 0.004 [-1.000, 1.000],	mean action:
896570/3500000:	episode: 409,	duration: 50.417s,	episode steps: 3998,	steps per second: 79,	episode reward: 2.000,	mean reward: 0.001 [-1.000, 1.000],	mean action:
900124/3500000:	episode: 410,	duration: 44.433s,	episode steps: 3554,	steps per second: 80,	episode reward: 7.000,	mean reward: 0.002 [-1.000, 1.000],	mean action:
903446/3500000:	episode: 411,	duration: 41.429s,	episode steps: 3322,	steps per second: 80,	episode reward: -4.000,	mean reward: -0.001 [-1.000, 1.000],	mean action:
907312/3500000:	episode: 412,	duration: 49.346s,	episode steps: 3866,	steps per second: 78,	episode reward: 8.000,	mean reward: 0.002 [-1.000, 1.000],	mean action:
910433/3500000:	episode: 413,	duration: 39.156s,	episode steps: 3121,	steps per second: 80,	episode reward: -12.000,	mean reward: -0.004 [-1.000, 1.000],	mean action:
913482/3500000:	episode: 414,	duration: 38.667s,	episode steps: 3049,	steps per second: 79,	episode reward: -11.000,	mean reward: -0.004 [-1.000, 1.000],	mean action:
917390/3500000:	episode: 415,	duration: 49.710s,	episode steps: 3908,	steps per second: 79,	episode reward: 6.000,	mean reward: 0.002 [-1.000, 1.000],	mean action:
920672/3500000:	episode: 416,	duration: 41.226s,	episode steps: 3282,	steps per second: 80,	episode reward: 6.000,	mean reward: 0.002 [-1.000, 1.000],	mean action:
924284/3500000:	episode: 417,	duration: 45.965s,	episode steps: 3612,	steps per second: 79,	episode reward: 4.000,	mean reward: 0.001 [-1.000, 1.000],	mean action:
927516/3500000:	episode: 418,	duration: 41.286s,	episode steps: 3232,	steps per second: 78,	episode reward: 6.000,	mean reward: 0.002 [-1.000, 1.000],	mean action:
929952/3500000:	episode: 419,	duration: 30.892s,	episode steps: 2436,	steps per second: 79,	episode reward: 13.000,	mean reward: 0.005 [-1.000, 1.000],	mean action:
933133/3500000:	episode: 420,	duration: 40.415s,	episode steps: 3181,	steps per second: 79,	episode reward: 7.000,	mean reward: 0.002 [-1.000, 1.000],	mean action:
936622/3500000:	episode: 421,	duration: 44.559s,	episode steps: 3489,	steps per second: 78,	episode reward: 4.000,	mean reward: 0.001 [-1.000, 1.000],	mean action:
940068/3500000:	episode: 422,	duration: 43.790s,	episode steps: 3446,	steps per second: 79,	episode reward: 5.000,	mean reward: 0.001 [-1.000, 1.000],	mean action:
944913/3500000:	episode: 423,	duration: 61.963s,	episode steps: 4845,	steps per second: 78,	episode reward: 1.000,	mean reward: 0.000 [-1.000, 1.000],	mean action:
949711/3500000:	episode: 424,	duration: 60.538s,	episode steps: 4798,	steps per second: 79,	episode reward: -1.000,	mean reward: -0.000 [-1.000, 1.000],	mean action:
953910/3500000:	episode: 425,	duration: 53.263s,	episode steps: 4199,	steps per second: 79,	episode reward: -1.000,	mean reward: -0.000 [-1.000, 1.000],	

995551/3500000:	episode:	436,	duration:	42.745s,	episode steps:	3435,	steps per second:	80,	episode reward:	8.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
998855/3500000:	episode:	437,	duration:	42.361s,	episode steps:	3304,	steps per second:	78,	episode reward:	6.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1002751/3500000:	episode:	438,	duration:	48.926s,	episode steps:	3896,	steps per second:	80,	episode reward:	6.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1006594/3500000:	episode:	439,	duration:	48.544s,	episode steps:	3843,	steps per second:	79,	episode reward:	3.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
1010558/3500000:	episode:	440,	duration:	50.273s,	episode steps:	3964,	steps per second:	79,	episode reward:	-1.000,	mean reward:	-0.000	[-1.000, 1.000],	mean action:
1014138/3500000:	episode:	441,	duration:	45.078s,	episode steps:	3580,	steps per second:	79,	episode reward:	4.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
1018319/3500000:	episode:	442,	duration:	52.580s,	episode steps:	4181,	steps per second:	80,	episode reward:	-3.000,	mean reward:	-0.001	[-1.000, 1.000],	mean action:
1022816/3500000:	episode:	443,	duration:	57.065s,	episode steps:	4497,	steps per second:	79,	episode reward:	3.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
1026142/3500000:	episode:	444,	duration:	41.987s,	episode steps:	3326,	steps per second:	79,	episode reward:	2.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
1029458/3500000:	episode:	445,	duration:	41.979s,	episode steps:	3316,	steps per second:	79,	episode reward:	5.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1031779/3500000:	episode:	446,	duration:	29.292s,	episode steps:	2321,	steps per second:	79,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1034979/3500000:	episode:	447,	duration:	40.043s,	episode steps:	3200,	steps per second:	80,	episode reward:	9.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
1038321/3500000:	episode:	448,	duration:	42.004s,	episode steps:	3342,	steps per second:	80,	episode reward:	-7.000,	mean reward:	-0.002	[-1.000, 1.000],	mean action:
1041623/3500000:	episode:	449,	duration:	41.824s,	episode steps:	3302,	steps per second:	79,	episode reward:	-2.000,	mean reward:	-0.001	[-1.000, 1.000],	mean action:
1044918/3500000:	episode:	450,	duration:	41.387s,	episode steps:	3295,	steps per second:	80,	episode reward:	-5.000,	mean reward:	-0.002	[-1.000, 1.000],	mean action:
1047945/3500000:	episode:	451,	duration:	38.741s,	episode steps:	3027,	steps per second:	78,	episode reward:	7.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1051397/3500000:	episode:	452,	duration:	43.900s,	episode steps:	3452,	steps per second:	79,	episode reward:	-3.000,	mean reward:	-0.001	[-1.000, 1.000],	mean action:
1054834/3500000:	episode:	453,	duration:	44.353s,	episode steps:	3437,	steps per second:	77,	episode reward:	1.000,	mean reward:	0.000	[-1.000, 1.000],	mean action:
1058529/3500000:	episode:	454,	duration:	47.411s,	episode steps:	3695,	steps per second:	78,	episode reward:	5.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
1062695/3500000:	episode:	455,	duration:	53.237s,	episode steps:	4166,	steps per second:	78,	episode reward:	5.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
1066160/3500000:	episode:	456,	duration:	43.792s,	episode steps:	3465,	steps per second:	79,	episode reward:	-5.000,	mean reward:	-0.001	[-1.000, 1.000],	mean action:
1069973/3500000:	episode:	457,	duration:	49.033s,	episode steps:	3813,	steps per second:	78,	episode reward:	5.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
1073485/3500000:	episode:	458,	duration:	44.856s,	episode steps:	3512,	steps per second:	78,	episode reward:	1.000,	mean reward:	0.000	[-1.000, 1.000],	mean action:
1076380/3500000:	episode:	459,	duration:	36.897s,	episode steps:	2895,	steps per second:	78,	episode reward:	11.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
1079769/3500000:	episode:	460,	duration:	42.832s,	episode steps:	3389,	steps per second:	79,	episode reward:	10.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
1082986/3500000:	episode:	461,	duration:	41.087s,	episode steps:	3217,	steps per second:	78,	episode reward:	12.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
1087183/3500000:	episode:	462,	duration:	53.466s,	episode steps:	4197,	steps per second:	78,	episode reward:	5.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
1089957/3500000:	episode:	463,	duration:	34.651s,	episode steps:	2774,	steps per second:	80,	episode reward:	13.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
1092348/3500000:	episode:	464,	duration:	29.564s,	episode steps:	2391,	steps per second:	81,	episode reward:	11.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
1095524/3500000:	episode:	465,	duration:	40.117s,	episode steps:	3176,	steps per second:	79,	episode reward:	5.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1098306/3500000:	episode:	466,	duration:	35.093s,	episode steps:	2782,	steps per second:	79,	episode reward:	10.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
1101108/3500000:	episode:	467,	duration:	34.810s,	episode steps:	2802,	steps per second:	80,	episode reward:	7.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1103953/3500000:	episode:	468,	duration:	36.225s,	episode steps:	2845,	steps per second:	79,	episode reward:	8.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
1106185/3500000:	episode:	469,	duration:	28.526s,	episode steps:	2232,	steps per second:	78,	episode reward:	15.000,	mean reward:	0.007	[-1.000, 1.000],	mean action:
1109399/3500000:	episode:	470,	duration:	41.247s,	episode steps:	3214,	steps per second:	78,	episode reward:	9.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
1112949/3500000:	episode:	471,	duration:	45.707s,	episode steps:	3550,	steps per second:	78,	episode reward:	5.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
1116799/3500000:	episode:	472,	duration:	49.260s,	episode steps:	3850,	steps per second:	78,	episode reward:	-2.000,	mean reward:	-0.001	[-1.000, 1.000],	mean action:
1119224/3500000:	episode:	473,	duration:	30.544s,	episode steps:	2425,	steps per second:	79,	episode reward:	13.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
1122163/3500000:	episode:	474,	duration:	36.531s,	episode steps:	2939,	steps per second:	80,	episode reward:	6.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1124489/3500000:	episode:	475,	duration:	28.773s,	episode steps:	2326,	steps per second:	81,	episode reward:	12.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
1127220/3500000:	episode:	476,	duration:	33.761s,	episode steps:	2731,	steps per second:	81,	episode reward:	-10.000,	mean reward:	-0.004	[-1.000, 1.000],	mean action:
1130296/3500000:	episode:	477,	duration:	38.617s,	episode steps:	3076,	steps per second:	80,	episode reward:	7.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1132651/3500000:	episode:	478,	duration:	29.208s,	episode steps:	2355,	steps per second:	81,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1135696/3500000:	episode:	479,	duration:	37.668s,	episode steps:	3045,	steps per second:	81,	episode reward:	-5.000,	mean reward:	-0.002	[-1.000, 1.000],	mean action:
1138645/3500000:	episode:	480,	duration:	36.539s,	episode steps:	2949,	steps per second:	81,	episode reward:	9.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
1141698/3500000:	episode:	481,	duration:	37.932s,	episode steps:	3053,	steps per second:	80,	episode reward:	-5.000,	mean reward:	-0.002	[-1.000, 1.000],	mean action:
1144783/3500000:	episode:	482,	duration:	39.707s,	episode steps:	3085,	steps per second:	78,	episode reward:	-7.000,	mean reward:	-0.002	[-1.000, 1.000],	mean action:
1148739/3500000:	episode:	483,	duration:	50.458s,	episode steps:	3956,	steps per second:	78,	episode reward:	-10.000,	mean reward:	-0.003	[-1.000, 1.000],	mean action:
1152241/3500000:	episode:	484,	duration:	45.222s,	episode steps:	3502,	steps per second:	77,	episode reward:	11.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
1155712/3500000:	episode:	485,	duration:	46.428s,	episode steps:	3471,	steps per second:	75,	episode reward:	2.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
1158854/3500000:	episode:	486,	duration:	40.921s,	episode steps:	3142,	steps per second:	77,	episode reward:	-8.000,	mean reward:	-0.003	[-1.000, 1.000],	mean action:
1162374/3500000:	episode:	487,	duration:	45.288s,	episode steps:	3520,	steps per second:	78,	episode reward:	1.000,	mean reward:	0.000	[-1.000, 1.000],	mean action:
1166426/3500000:	episode:	488,	duration:	51.785s,	episode steps:	4052,	steps per second:	78,	episode reward:	1.000,	mean reward:	0.000	[-1.000, 1.000],	mean action:
1169579/3500000:	episode:	489,	duration:	39.561s,	episode steps:	3153,	steps per second:	80,	episode reward:	-6.000,	mean reward:	-0.002	[-1.000, 1.000],	mean action:
1173409/3500000:	episode:	490,	duration:	49.028s,	episode steps:	3830,	steps per second:	78,	episode reward:	-5.000,	mean reward:	-0.001	[-1.000, 1.000],	mean action:
1176453/3500000:	episode:	491,	duration:	38.803s,	episode steps:	3044,	steps per second:	78,	episode reward:	6.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1179931/3500000:	episode:	492,	duration:	44.361s,	episode steps:	3478,	steps per second:	78,	episode reward:	-6.000,	mean reward:	-0.002	[-1.000, 1.000],	mean action:
1184636/3500000:	episode:	493,	duration:	60.159s,	episode steps:	4705,	steps per second:	78,	episode reward:	-2.000,	mean reward:	-0.000	[-1.000, 1.000],	mean action:
1188230/3500000:	episode:	494,	duration:	46.025s,	episode steps:	3594,	steps per second:	78,	episode reward:	8.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1192135/3500000:	episode:	495,	duration:	49.220s,	episode steps:	3905,	steps per second:	79,	episode reward:	-4.000,	mean reward:	-0.001	[-1.000, 1.000],	mean action:
1196363/3500000:	episode:	496,	duration:	53.183s,	episode steps:	4228,	steps per second:	79,	episode reward:	-1.000,	mean reward:	-0.000	[-1.000, 1.000],	mean action:
1199583/3500000:	episode:	497,	duration:	40.462s,	episode steps:	3220,	steps per second:	80,	episode reward:	9.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
1203142/3500000:	episode:	498,	duration:	44.447s,	episode steps:	3559,	steps per second:	80,	episode reward:	-5.000,	mean reward:	-0.001	[-1.000, 1.000],	mean action:
1207784/3500000:	episode:	499,	duration:	58.218s,	episode steps:	4642,	steps per second:	80,	episode reward:	3.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
1210292/3500000:	episode:	500,	duration:	31.193s,	episode steps:	2508,	steps per second:	80,	episode reward:	11.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
1213354/3500000:	episode:	501,	duration:	37.573s,	episode steps:	3062,	steps per second:	81,	episode reward:	3.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
1217208/3500000:	episode:	502,	duration:	47.760s,	episode steps:	3854,	steps per second:	81,	episode reward:	8.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1221552/3500000:	episode:	503,	duration:	53.997s,	episode steps:	4344,	steps per second:	80,	episode reward:	-4.000,	mean reward:	-0.001	[-1.000, 1.000],	mean action:
1224775/3500000:	episode:	504,	duration:	40.266s,	episode steps:	3223,	steps per second:	80,	episode reward:	-1.000,	mean reward:	-0.000	[-1.000, 1.000],	mean action:
1228186/3500000:	episode:	505,	duration:	42.271s,	episode steps:	3411,	steps per second:	81,	episode reward:	6.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1231878/3500000:	episode:	506,	duration:	45.959s,	episode steps:	3692,	steps per second:	80,	episode reward:	2.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
1235359/3500000:	episode:	507,	duration:	43.101s,	episode steps:	3481,	steps per second:	81,	episode reward:	4.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
1238283/3500000:	episode:	508,	duration:	36.593s,	episode steps:	2924,	steps per second:	80,	episode reward:	8.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
1242079/3500000:	episode:	509,	duration:	46.779s,	episode steps:	3796,	steps per second:	81,	episode reward:	7.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1246066/3500000:	episode:	510,	duration:	49.170s,	episode steps:	3987,	steps per second:	81,	episode reward:	-6.000,	mean reward:	-0.002	[-1.000, 1.000],	mean action:
1250072/3500000:	episode:	511,	duration:	50.058s,	episode steps:	4006,	steps per second:	80,	episode reward:	5.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
1254213/3500000:	episode:	512,	duration:	51.587s,	episode steps:	4141,	steps per second:	80,	episode reward:	7.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1257621/3500000:	episode:	513,	duration:	43.084s,	episode steps:	3408,	steps per second:	79,	episode reward:	8.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1261737/3500000:	episode:	514,	duration:	51.803s,	episode steps:	4116,	steps per second:	79,	episode reward:	3.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
1266447/3500000:	episode:	515,	duration:	58.476s,	episode steps:	4710,	steps per second:	81,	episode reward:	6.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
1269792/3500000:	episode:	516,	duration:	41.744s,	episode steps:	3345,	steps per second:	80,	episode reward:	8.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1273512/3500000:	episode:	517,	duration:	46.340s,	episode steps:	3720,	steps per second:	80,	episode reward:	-5.000,	mean reward:	-0.001	[-1.000, 1.000],	mean action:
1277142/3500000:	episode:	518,	duration:	45.131s,	episode steps:	3630,	steps per second:	80,	episode reward:	3.000,	mean reward:	0.001	[-1.000, 1.000],	

```

1369693/3500000: episode: 545, duration: 52.077s, episode steps: 4215, steps per second: 81, episode reward: 2.000, mean reward: 0.000 [-1.000, 1.000], mean action:
1373432/3500000: episode: 546, duration: 46.253s, episode steps: 3739, steps per second: 81, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean action:
1376528/3500000: episode: 547, duration: 39.000s, episode steps: 3096, steps per second: 79, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1379471/3500000: episode: 548, duration: 36.777s, episode steps: 2943, steps per second: 80, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1382437/3500000: episode: 549, duration: 37.077s, episode steps: 2966, steps per second: 80, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1385665/3500000: episode: 550, duration: 41.062s, episode steps: 3228, steps per second: 79, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1388844/3500000: episode: 551, duration: 39.907s, episode steps: 3179, steps per second: 80, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1393139/3500000: episode: 552, duration: 53.377s, episode steps: 4295, steps per second: 80, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean action:
1396553/3500000: episode: 553, duration: 42.496s, episode steps: 3414, steps per second: 80, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1399868/3500000: episode: 554, duration: 41.448s, episode steps: 3315, steps per second: 80, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1402035/3500000: episode: 555, duration: 27.624s, episode steps: 2167, steps per second: 78, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean action:
1405820/3500000: episode: 556, duration: 47.483s, episode steps: 3785, steps per second: 80, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1409193/3500000: episode: 557, duration: 41.820s, episode steps: 3373, steps per second: 81, episode reward: 11.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1412022/3500000: episode: 558, duration: 35.056s, episode steps: 2829, steps per second: 81, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1415500/3500000: episode: 559, duration: 42.900s, episode steps: 3478, steps per second: 81, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1420455/3500000: episode: 560, duration: 62.532s, episode steps: 4955, steps per second: 79, episode reward: 2.000, mean reward: 0.000 [-1.000, 1.000], mean action:
1424162/3500000: episode: 561, duration: 47.074s, episode steps: 3707, steps per second: 79, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1427423/3500000: episode: 562, duration: 40.736s, episode steps: 3261, steps per second: 80, episode reward: 11.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1432037/3500000: episode: 563, duration: 57.741s, episode steps: 4614, steps per second: 80, episode reward: 2.000, mean reward: 0.000 [-1.000, 1.000], mean action:
1435466/3500000: episode: 564, duration: 42.908s, episode steps: 3429, steps per second: 80, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1438899/3500000: episode: 565, duration: 42.869s, episode steps: 3433, steps per second: 80, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1442120/3500000: episode: 566, duration: 40.687s, episode steps: 3221, steps per second: 79, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1444934/3500000: episode: 567, duration: 35.474s, episode steps: 2814, steps per second: 79, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1448541/3500000: episode: 568, duration: 44.982s, episode steps: 3607, steps per second: 80, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
1450866/3500000: episode: 569, duration: 29.447s, episode steps: 2325, steps per second: 79, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1453928/3500000: episode: 570, duration: 38.549s, episode steps: 3062, steps per second: 79, episode reward: 13.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1456893/3500000: episode: 571, duration: 37.278s, episode steps: 2965, steps per second: 80, episode reward: 15.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1460567/3500000: episode: 572, duration: 45.863s, episode steps: 3674, steps per second: 80, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1463509/3500000: episode: 573, duration: 36.804s, episode steps: 2942, steps per second: 80, episode reward: 13.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1466803/3500000: episode: 574, duration: 40.970s, episode steps: 3294, steps per second: 80, episode reward: 11.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1470464/3500000: episode: 575, duration: 45.677s, episode steps: 3661, steps per second: 80, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1473788/3500000: episode: 576, duration: 41.880s, episode steps: 3324, steps per second: 79, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1477090/3500000: episode: 577, duration: 41.121s, episode steps: 3302, steps per second: 80, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1481029/3500000: episode: 578, duration: 49.031s, episode steps: 3939, steps per second: 80, episode reward: 9.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1483998/3500000: episode: 579, duration: 37.569s, episode steps: 2969, steps per second: 79, episode reward: 13.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1486558/3500000: episode: 580, duration: 31.427s, episode steps: 2560, steps per second: 81, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1489168/3500000: episode: 581, duration: 32.078s, episode steps: 2610, steps per second: 81, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1491768/3500000: episode: 582, duration: 32.503s, episode steps: 2600, steps per second: 80, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1494556/3500000: episode: 583, duration: 34.641s, episode steps: 2788, steps per second: 80, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1497422/3500000: episode: 584, duration: 35.419s, episode steps: 2866, steps per second: 81, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1499789/3500000: episode: 585, duration: 29.223s, episode steps: 2367, steps per second: 81, episode reward: 18.000, mean reward: 0.008 [-1.000, 1.000], mean action:
1502345/3500000: episode: 586, duration: 31.950s, episode steps: 2556, steps per second: 80, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1504705/3500000: episode: 587, duration: 29.228s, episode steps: 2360, steps per second: 81, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1509047/3500000: episode: 588, duration: 54.147s, episode steps: 4342, steps per second: 80, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean action:
1511582/3500000: episode: 589, duration: 31.401s, episode steps: 2535, steps per second: 81, episode reward: 16.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1514841/3500000: episode: 590, duration: 40.379s, episode steps: 3259, steps per second: 81, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1517356/3500000: episode: 591, duration: 30.638s, episode steps: 2515, steps per second: 82, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1519897/3500000: episode: 592, duration: 31.366s, episode steps: 2541, steps per second: 81, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean action:
1522838/3500000: episode: 593, duration: 36.554s, episode steps: 2941, steps per second: 80, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1525846/3500000: episode: 594, duration: 37.020s, episode steps: 3008, steps per second: 81, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1528768/3500000: episode: 595, duration: 35.490s, episode steps: 2922, steps per second: 82, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1531949/3500000: episode: 596, duration: 39.419s, episode steps: 3181, steps per second: 81, episode reward: 11.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1534972/3500000: episode: 597, duration: 37.443s, episode steps: 3023, steps per second: 81, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1538768/3500000: episode: 598, duration: 47.328s, episode steps: 3796, steps per second: 80, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1541905/3500000: episode: 599, duration: 38.889s, episode steps: 3137, steps per second: 81, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1545313/3500000: episode: 600, duration: 42.100s, episode steps: 3408, steps per second: 81, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1548426/3500000: episode: 601, duration: 38.387s, episode steps: 3113, steps per second: 81, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1551874/3500000: episode: 602, duration: 42.251s, episode steps: 3448, steps per second: 82, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1555808/3500000: episode: 603, duration: 47.958s, episode steps: 3934, steps per second: 82, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1558990/3500000: episode: 604, duration: 38.786s, episode steps: 3182, steps per second: 82, episode reward: 13.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1563792/3500000: episode: 605, duration: 58.458s, episode steps: 4802, steps per second: 82, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
1567095/3500000: episode: 606, duration: 40.183s, episode steps: 3303, steps per second: 82, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1570434/3500000: episode: 607, duration: 41.306s, episode steps: 3339, steps per second: 81, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1573254/3500000: episode: 608, duration: 35.133s, episode steps: 2820, steps per second: 80, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1576468/3500000: episode: 609, duration: 39.765s, episode steps: 3214, steps per second: 81, episode reward: 14.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1578844/3500000: episode: 610, duration: 29.396s, episode steps: 2376, steps per second: 81, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean action:
1582060/3500000: episode: 611, duration: 40.474s, episode steps: 3216, steps per second: 79, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1586655/3500000: episode: 612, duration: 56.532s, episode steps: 4595, steps per second: 81, episode reward: 2.000, mean reward: 0.000 [-1.000, 1.000], mean action:
1589510/3500000: episode: 613, duration: 35.186s, episode steps: 2855, steps per second: 81, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1593041/3500000: episode: 614, duration: 43.821s, episode steps: 3531, steps per second: 81, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1596692/3500000: episode: 615, duration: 45.437s, episode steps: 3651, steps per second: 80, episode reward: 9.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1599708/3500000: episode: 616, duration: 37.621s, episode steps: 3016, steps per second: 80, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1602069/3500000: episode: 617, duration: 30.549s, episode steps: 2361, steps per second: 77, episode reward: 13.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1605092/3500000: episode: 618, duration: 43.466s, episode steps: 3433, steps per second: 79, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1608665/3500000: episode: 619, duration: 39.951s, episode steps: 3163, steps per second: 79, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1611963/3500000: episode: 620, duration: 41.364s, episode steps: 3298, steps per second: 80, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1615177/3500000: episode: 621, duration: 40.545s, episode steps: 3214, steps per second: 79, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1618004/3500000: episode: 622, duration: 35.163s, episode steps: 2827, steps per second: 80, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1620869/3500000: episode: 623, duration: 35.697s, episode steps: 2865, steps per second: 80, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1622964/3500000: episode: 624, duration: 25.712s, episode steps: 2095, steps per second: 81, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean action:
1626189/3500000: episode: 625, duration: 40.808s, episode steps: 3225, steps per second: 79, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1628980/3500000: episode: 626, duration: 34.457s, episode steps: 2791, steps per second: 81, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1632307/3500000: episode: 627, duration: 41.693s, episode steps: 3327, steps per second: 80, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1634206/3500000: episode: 628, duration: 23.624s, episode steps: 1899, steps per second: 80, episode reward: 20.000, mean reward: 0.011 [-1.000, 1.000], mean action:
1637472/3500000: episode: 629, duration: 41.164s, episode steps: 3266, steps per second: 79, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1640451/3500000: episode: 630, duration: 37.688s, episode steps: 2979, steps per second: 79, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1643150/3500000: episode: 631, duration: 33.763s, episode steps: 2699, steps per second: 80, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1646057/3500000: episode: 632, duration: 36.141s, episode steps: 2907, steps per second: 80, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1648527/3500000: episode: 633, duration: 31.036s, episode steps: 2470, steps per second: 80, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1651275/3500000: episode: 634, duration: 34.600s, episode steps: 2748, steps per second: 79, episode reward: 16.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1654110/3500000: episode: 635, duration: 35.829s, episode steps: 2835, steps per second: 79, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1656974/3500000: episode: 636, duration: 35.725s, episode steps: 2864, steps per second: 80, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1659305/3500000: episode: 637, duration: 29.251s, episode steps: 2331, steps per second: 80, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1662198/3500000: episode: 638, duration: 35.887s, episode steps: 2893, steps per second: 81, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1665648/3500000: episode: 639, duration: 43.585s, episode steps: 3450, steps per second: 79, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1668752/3500000: episode: 640, duration: 40.071s, episode steps: 3104, steps per second: 77, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1671184/3500000: episode: 641, duration: 30.390s, episode steps: 2432, steps per second: 80, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1674183/3500000: episode: 642, duration: 37.864s, episode steps: 2999, steps per second: 79, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1676609/3500000: episode: 643, duration: 29.976s, episode steps: 2426, steps per second: 81, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1678679/3500000: episode: 644, duration: 25.948s, episode steps: 2070, steps per second: 80, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean action:
1681072/3500000: episode: 645, duration: 29.896s, episode steps: 2393, steps per second: 80, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1683567/3500000: episode: 646, duration: 30.899s, episode steps: 2495, steps per second: 81, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean action:
1686856/3500000: episode: 647, duration: 41.744s, episode steps: 3289, steps per second: 79, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1690021/3500000: episode: 648, duration: 39.910s, episode steps: 3165, steps per second: 79, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1692778/3500000: episode: 649, duration: 35.606s, episode steps: 2757, steps per second: 77, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1695838/3500000: episode: 650, duration: 39.891s, episode steps: 3060, steps per second: 77, episode reward: 13.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1698444/3500000: episode: 651, duration: 33.312s, episode steps: 2606, steps per second: 78, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1701098/3500000: episode: 652, duration: 33.414s, episode steps: 2654, steps per second: 79, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action:
1703826/3500000: episode: 653, duration: 34.763s, episode steps: 2728, steps per second: 78, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action:

```

1706021/3500000:	episode:	654,	duration:	27.856s,	episode steps:	2195,	steps per second:	79,	episode reward:	16.000,	mean reward:	0.007	[-1.000, 1.000],	mean action:
1708522/3500000:	episode:	655,	duration:	31.839s,	episode steps:	2501,	steps per second:	79,	episode reward:	15.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1710908/3500000:	episode:	656,	duration:	29.673s,	episode steps:	2386,	steps per second:	80,	episode reward:	15.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1713097/3500000:	episode:	657,	duration:	27.864s,	episode steps:	2189,	steps per second:	79,	episode reward:	16.000,	mean reward:	0.007	[-1.000, 1.000],	mean action:
1715504/3500000:	episode:	658,	duration:	30.605s,	episode steps:	2407,	steps per second:	79,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1717660/3500000:	episode:	659,	duration:	27.019s,	episode steps:	2156,	steps per second:	80,	episode reward:	17.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
1719949/3500000:	episode:	660,	duration:	29.772s,	episode steps:	2289,	steps per second:	77,	episode reward:	16.000,	mean reward:	0.007	[-1.000, 1.000],	mean action:
1722139/3500000:	episode:	661,	duration:	2957.398s,	episode steps:	2190,	steps per second:	1,	episode reward:	18.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
1724003/3500000:	episode:	662,	duration:	22.559s,	episode steps:	1861,	steps per second:	82,	episode reward:	20.000,	mean reward:	0.011	[-1.000, 1.000],	mean action:
1727016/3500000:	episode:	663,	duration:	37.523s,	episode steps:	3016,	steps per second:	80,	episode reward:	11.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
1729519/3500000:	episode:	664,	duration:	30.352s,	episode steps:	2503,	steps per second:	82,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1731684/3500000:	episode:	665,	duration:	25.486s,	episode steps:	2165,	steps per second:	85,	episode reward:	17.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
1734250/3500000:	episode:	666,	duration:	30.687s,	episode steps:	2566,	steps per second:	84,	episode reward:	14.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
1736198/3500000:	episode:	667,	duration:	23.346s,	episode steps:	1948,	steps per second:	83,	episode reward:	20.000,	mean reward:	0.010	[-1.000, 1.000],	mean action:
1738321/3500000:	episode:	668,	duration:	25.322s,	episode steps:	2123,	steps per second:	84,	episode reward:	16.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
1740798/3500000:	episode:	669,	duration:	30.472s,	episode steps:	2477,	steps per second:	81,	episode reward:	17.000,	mean reward:	0.007	[-1.000, 1.000],	mean action:
1743176/3500000:	episode:	670,	duration:	28.246s,	episode steps:	2378,	steps per second:	84,	episode reward:	18.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
1746153/3500000:	episode:	671,	duration:	35.956s,	episode steps:	2977,	steps per second:	83,	episode reward:	8.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
1748528/3500000:	episode:	672,	duration:	28.275s,	episode steps:	2375,	steps per second:	84,	episode reward:	13.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
1751653/3500000:	episode:	673,	duration:	37.828s,	episode steps:	3125,	steps per second:	83,	episode reward:	10.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
1753909/3500000:	episode:	674,	duration:	27.402s,	episode steps:	2256,	steps per second:	82,	episode reward:	15.000,	mean reward:	0.007	[-1.000, 1.000],	mean action:
1757519/3500000:	episode:	675,	duration:	44.538s,	episode steps:	3610,	steps per second:	81,	episode reward:	9.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1760430/3500000:	episode:	676,	duration:	35.278s,	episode steps:	2911,	steps per second:	83,	episode reward:	10.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
1763025/3500000:	episode:	677,	duration:	31.588s,	episode steps:	2595,	steps per second:	82,	episode reward:	13.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
1765461/3500000:	episode:	678,	duration:	29.944s,	episode steps:	2436,	steps per second:	81,	episode reward:	15.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1767499/3500000:	episode:	679,	duration:	24.383s,	episode steps:	2038,	steps per second:	84,	episode reward:	19.000,	mean reward:	0.009	[-1.000, 1.000],	mean action:
1769709/3500000:	episode:	680,	duration:	27.107s,	episode steps:	2210,	steps per second:	82,	episode reward:	17.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
1772143/3500000:	episode:	681,	duration:	33.237s,	episode steps:	2704,	steps per second:	81,	episode reward:	14.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
1776183/3500000:	episode:	682,	duration:	44.619s,	episode steps:	3770,	steps per second:	84,	episode reward:	4.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
1779417/3500000:	episode:	683,	duration:	38.164s,	episode steps:	3234,	steps per second:	85,	episode reward:	12.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
1781680/3500000:	episode:	684,	duration:	26.658s,	episode steps:	2263,	steps per second:	85,	episode reward:	17.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
1784048/3500000:	episode:	685,	duration:	28.005s,	episode steps:	2368,	steps per second:	85,	episode reward:	16.000,	mean reward:	0.007	[-1.000, 1.000],	mean action:
1787455/3500000:	episode:	686,	duration:	40.126s,	episode steps:	3407,	steps per second:	85,	episode reward:	7.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1789183/3500000:	episode:	687,	duration:	20.550s,	episode steps:	1728,	steps per second:	84,	episode reward:	19.000,	mean reward:	0.011	[-1.000, 1.000],	mean action:
1791484/3500000:	episode:	688,	duration:	27.317s,	episode steps:	2301,	steps per second:	84,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1793856/3500000:	episode:	689,	duration:	28.183s,	episode steps:	2372,	steps per second:	84,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1796257/3500000:	episode:	690,	duration:	28.296s,	episode steps:	2401,	steps per second:	85,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1799501/3500000:	episode:	691,	duration:	38.234s,	episode steps:	3244,	steps per second:	85,	episode reward:	12.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
1801407/3500000:	episode:	692,	duration:	22.463s,	episode steps:	1906,	steps per second:	85,	episode reward:	20.000,	mean reward:	0.010	[-1.000, 1.000],	mean action:
1803612/3500000:	episode:	693,	duration:	25.865s,	episode steps:	2205,	steps per second:	85,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1805742/3500000:	episode:	694,	duration:	25.079s,	episode steps:	2130,	steps per second:	85,	episode reward:	18.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
1809207/3500000:	episode:	695,	duration:	40.647s,	episode steps:	3465,	steps per second:	85,	episode reward:	10.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
1812224/3500000:	episode:	696,	duration:	35.458s,	episode steps:	3017,	steps per second:	85,	episode reward:	8.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
1815895/3500000:	episode:	697,	duration:	43.044s,	episode steps:	3671,	steps per second:	85,	episode reward:	9.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1818892/3500000:	episode:	698,	duration:	35.176s,	episode steps:	2997,	steps per second:	85,	episode reward:	11.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
1821226/3500000:	episode:	699,	duration:	27.445s,	episode steps:	2334,	steps per second:	85,	episode reward:	16.000,	mean reward:	0.007	[-1.000, 1.000],	mean action:
1824078/3500000:	episode:	700,	duration:	33.458s,	episode steps:	2852,	steps per second:	85,	episode reward:	15.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
1826953/3500000:	episode:	701,	duration:	33.797s,	episode steps:	2875,	steps per second:	85,	episode reward:	9.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
1830705/3500000:	episode:	702,	duration:	44.086s,	episode steps:	3752,	steps per second:	85,	episode reward:	9.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1833068/3500000:	episode:	703,	duration:	27.677s,	episode steps:	2363,	steps per second:	85,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1835953/3500000:	episode:	704,	duration:	33.866s,	episode steps:	2885,	steps per second:	85,	episode reward:	15.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
1839155/3500000:	episode:	705,	duration:	37.476s,	episode steps:	3202,	steps per second:	85,	episode reward:	11.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
1841657/3500000:	episode:	706,	duration:	29.422s,	episode steps:	2502,	steps per second:	85,	episode reward:	13.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
1844210/3500000:	episode:	707,	duration:	29.857s,	episode steps:	2553,	steps per second:	86,	episode reward:	15.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1846785/3500000:	episode:	708,	duration:	30.958s,	episode steps:	2575,	steps per second:	83,	episode reward:	14.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
1849357/3500000:	episode:	709,	duration:	30.256s,	episode steps:	2572,	steps per second:	85,	episode reward:	17.000,	mean reward:	0.007	[-1.000, 1.000],	mean action:
1851462/3500000:	episode:	710,	duration:	24.809s,	episode steps:	2105,	steps per second:	85,	episode reward:	20.000,	mean reward:	0.010	[-1.000, 1.000],	mean action:
1854310/3500000:	episode:	711,	duration:	33.297s,	episode steps:	2848,	steps per second:	86,	episode reward:	13.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
1856798/3500000:	episode:	712,	duration:	29.228s,	episode steps:	2488,	steps per second:	85,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1859502/3500000:	episode:	713,	duration:	31.645s,	episode steps:	2704,	steps per second:	85,	episode reward:	14.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
1861899/3500000:	episode:	714,	duration:	28.343s,	episode steps:	2397,	steps per second:	85,	episode reward:	15.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1864298/3500000:	episode:	715,	duration:	28.145s,	episode steps:	2399,	steps per second:	85,	episode reward:	15.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1866818/3500000:	episode:	716,	duration:	29.651s,	episode steps:	2520,	steps per second:	85,	episode reward:	13.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
1869037/3500000:	episode:	717,	duration:	26.041s,	episode steps:	2219,	steps per second:	85,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1872296/3500000:	episode:	718,	duration:	38.444s,	episode steps:	3259,	steps per second:	85,	episode reward:	10.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
1874913/3500000:	episode:	719,	duration:	30.703s,	episode steps:	2617,	steps per second:	85,	episode reward:	12.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
1877172/3500000:	episode:	720,	duration:	26.504s,	episode steps:	2259,	steps per second:	85,	episode reward:	18.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
1879609/3500000:	episode:	721,	duration:	28.500s,	episode steps:	2437,	steps per second:	86,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1882799/3500000:	episode:	722,	duration:	37.433s,	episode steps:	3190,	steps per second:	85,	episode reward:	9.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
1885049/3500000:	episode:	723,	duration:	26.398s,	episode steps:	2250,	steps per second:	85,	episode reward:	15.000,	mean reward:	0.007	[-1.000, 1.000],	mean action:
1888837/3500000:	episode:	724,	duration:	44.868s,	episode steps:	3788,	steps per second:	84,	episode reward:	8.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1891654/3500000:	episode:	725,	duration:	33.060s,	episode steps:	2817,	steps per second:	85,	episode reward:	12.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
1894175/3500000:	episode:	726,	duration:	29.486s,	episode steps:	2521,	steps per second:	85,	episode reward:	11.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
1897202/3500000:	episode:	727,	duration:	35.527s,	episode steps:	3027,	steps per second:	85,	episode reward:	13.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
1900448/3500000:	episode:	728,	duration:	37.989s,	episode steps:	3246,	steps per second:	85,	episode reward:	11.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
1903355/3500000:	episode:	729,	duration:	33.924s,	episode steps:	2907,	steps per second:	86,	episode reward:	13.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
1905726/3500000:	episode:	730,	duration:	27.690s,	episode steps:	2371,	steps per second:	86,	episode reward:	16.000,	mean reward:	0.007	[-1.000, 1.000],	mean action:
1908305/3500000:	episode:	731,	duration:	30.064s,	episode steps:	2579,	steps per second:	86,	episode reward:	11.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
1911043/3500000:	episode:	732,	duration:	32.013s,	episode steps:	2738,	steps per second:	86,	episode reward:	10.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
1913783/3500000:	episode:	733,	duration:	32.081s,	episode steps:	2740,	steps per second:	85,	episode reward:	16.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1916665/3500000:	episode:	734,	duration:	33.691s,	episode steps:	2882,	steps per second:	86,	episode reward:	7.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
1919002/3500000:	episode:	735,	duration:	27.338s,	episode steps:	2337,	steps per second:	85,	episode reward:	15.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
1921794/3500000:	episode:	736,	duration:	32.632s,	episode steps:	2792,	steps per second:	8						

1992273/3500000:	episode: 763,	duration: 25.235s,	episode steps: 1976,	steps per second: 85,	episode reward: 20.000,	mean reward: 0.010	[-1.000, 1.000],	mean action
1990227/3500000:	episode: 764,	duration: 29.885s,	episode steps: 2548,	steps per second: 85,	episode reward: 15.000,	mean reward: 0.006	[-1.000, 1.000],	mean action
1992533/3500000:	episode: 765,	duration: 26.996s,	episode steps: 2306,	steps per second: 85,	episode reward: 18.000,	mean reward: 0.008	[-1.000, 1.000],	mean action
1995527/3500000:	episode: 766,	duration: 32.801s,	episode steps: 2794,	steps per second: 85,	episode reward: 10.000,	mean reward: 0.004	[-1.000, 1.000],	mean action
1997442/3500000:	episode: 767,	duration: 24.730s,	episode steps: 2115,	steps per second: 86,	episode reward: 16.000,	mean reward: 0.008	[-1.000, 1.000],	mean action
1999344/3500000:	episode: 768,	duration: 23.433s,	episode steps: 1992,	steps per second: 85,	episode reward: 18.000,	mean reward: 0.009	[-1.000, 1.000],	mean action
2001451/3500000:	episode: 769,	duration: 23.734s,	episode steps: 2017,	steps per second: 85,	episode reward: 16.000,	mean reward: 0.008	[-1.000, 1.000],	mean action
2005702/3500000:	episode: 770,	duration: 50.035s,	episode steps: 4251,	steps per second: 85,	episode reward: 5.000,	mean reward: 0.001	[-1.000, 1.000],	mean action
2008685/3500000:	episode: 771,	duration: 34.898s,	episode steps: 2983,	steps per second: 85,	episode reward: 10.000,	mean reward: 0.003	[-1.000, 1.000],	mean action
2011730/3500000:	episode: 772,	duration: 35.490s,	episode steps: 3045,	steps per second: 86,	episode reward: 15.000,	mean reward: 0.005	[-1.000, 1.000],	mean action
2014833/3500000:	episode: 773,	duration: 29.842s,	episode steps: 2553,	steps per second: 86,	episode reward: 16.000,	mean reward: 0.006	[-1.000, 1.000],	mean action
2016909/3500000:	episode: 774,	duration: 30.620s,	episode steps: 2626,	steps per second: 86,	episode reward: 13.000,	mean reward: 0.005	[-1.000, 1.000],	mean action
2019044/3500000:	episode: 775,	duration: 25.063s,	episode steps: 2135,	steps per second: 85,	episode reward: 18.000,	mean reward: 0.008	[-1.000, 1.000],	mean action
2021452/3500000:	episode: 776,	duration: 28.233s,	episode steps: 2408,	steps per second: 85,	episode reward: 16.000,	mean reward: 0.007	[-1.000, 1.000],	mean action
2023660/3500000:	episode: 777,	duration: 25.991s,	episode steps: 2208,	steps per second: 85,	episode reward: 19.000,	mean reward: 0.009	[-1.000, 1.000],	mean action
2027832/3500000:	episode: 778,	duration: 49.131s,	episode steps: 4172,	steps per second: 85,	episode reward: 3.000,	mean reward: 0.001	[-1.000, 1.000],	mean action
2029842/3500000:	episode: 779,	duration: 23.620s,	episode steps: 2010,	steps per second: 85,	episode reward: 17.000,	mean reward: 0.008	[-1.000, 1.000],	mean action
2032394/3500000:	episode: 780,	duration: 29.991s,	episode steps: 2552,	steps per second: 85,	episode reward: 14.000,	mean reward: 0.005	[-1.000, 1.000],	mean action
2034780/3500000:	episode: 781,	duration: 27.884s,	episode steps: 2386,	steps per second: 86,	episode reward: 15.000,	mean reward: 0.006	[-1.000, 1.000],	mean action
2037486/3500000:	episode: 782,	duration: 31.572s,	episode steps: 2706,	steps per second: 86,	episode reward: 16.000,	mean reward: 0.006	[-1.000, 1.000],	mean action
2039553/3500000:	episode: 783,	duration: 24.132s,	episode steps: 2067,	steps per second: 86,	episode reward: 18.000,	mean reward: 0.009	[-1.000, 1.000],	mean action
2042967/3500000:	episode: 784,	duration: 40.569s,	episode steps: 3414,	steps per second: 84,	episode reward: 7.000,	mean reward: 0.002	[-1.000, 1.000],	mean action
2045602/3500000:	episode: 785,	duration: 31.111s,	episode steps: 2635,	steps per second: 85,	episode reward: 11.000,	mean reward: 0.004	[-1.000, 1.000],	mean action
2048291/3500000:	episode: 786,	duration: 31.387s,	episode steps: 2689,	steps per second: 86,	episode reward: 11.000,	mean reward: 0.004	[-1.000, 1.000],	mean action
0511277/3500000:	episode: 787,	duration: 33.405s,	episode steps: 2836,	steps per second: 85,	episode reward: 12.000,	mean reward: 0.004	[-1.000, 1.000],	mean action
0545477/3500000:	episode: 788,	duration: 41.526s,	episode steps: 3450,	steps per second: 83,	episode reward: 7.000,	mean reward: 0.002	[-1.000, 1.000],	mean action
0572255/3500000:	episode: 789,	duration: 31.398s,	episode steps: 2678,	steps per second: 85,	episode reward: 11.000,	mean reward: 0.004	[-1.000, 1.000],	mean action
0600759/3500000:	episode: 790,	duration: 40.985s,	episode steps: 3504,	steps per second: 85,	episode reward: 10.000,	mean reward: 0.003	[-1.000, 1.000],	mean action
0603726/3500000:	episode: 791,	duration: 34.578s,	episode steps: 2967,	steps per second: 86,	episode reward: 10.000,	mean reward: 0.003	[-1.000, 1.000],	mean action
0605824/3500000:	episode: 792,	duration: 25.817s,	episode steps: 2198,	steps per second: 85,	episode reward: 16.000,	mean reward: 0.007	[-1.000, 1.000],	mean action
2068944/3500000:	episode: 793,	duration: 34.021s,	episode steps: 2920,	steps per second: 86,	episode reward: 12.000,	mean reward: 0.004	[-1.000, 1.000],	mean action
2070778/3500000:	episode: 794,	duration: 22.670s,	episode steps: 1934,	steps per second:				

2274706/3500000: episode: 872, duration: 40.301s, episode steps: 3423, steps per second: 85, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action: 2278826/3500000: episode: 873, duration: 48.474s, episode steps: 4120, steps per second: 85, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action: 2280978/3500000: episode: 874, duration: 25.339s, episode steps: 2152, steps per second: 85, episode reward: 18.000, mean reward: 0.008 [-1.000, 1.000], mean action: 2284851/3500000: episode: 875, duration: 45.378s, episode steps: 3873, steps per second: 85, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action: 2287397/3500000: episode: 876, duration: 29.894s, episode steps: 2546, steps per second: 85, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean action: 2289785/3500000: episode: 877, duration: 28.097s, episode steps: 2388, steps per second: 85, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2291917/3500000: episode: 878, duration: 25.078s, episode steps: 2132, steps per second: 85, episode reward: 18.000, mean reward: 0.008 [-1.000, 1.000], mean action: 2294224/3500000: episode: 879, duration: 27.088s, episode steps: 2307, steps per second: 85, episode reward: 13.000, mean reward: 0.006 [-1.000, 1.000], mean action: 2296517/3500000: episode: 880, duration: 27.050s, episode steps: 2293, steps per second: 85, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2299993/3500000: episode: 881, duration: 40.847s, episode steps: 3476, steps per second: 85, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean action: 2302754/3500000: episode: 882, duration: 32.363s, episode steps: 2761, steps per second: 85, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2305830/3500000: episode: 883, duration: 36.101s, episode steps: 3076, steps per second: 85, episode reward: 13.000, mean reward: 0.004 [-1.000, 1.000], mean action: 2307970/3500000: episode: 884, duration: 25.192s, episode steps: 2140, steps per second: 85, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2311280/3500000: episode: 885, duration: 38.864s, episode steps: 3310, steps per second: 85, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action: 2315473/3500000: episode: 886, duration: 49.424s, episode steps: 4193, steps per second: 85, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action: 2317612/3500000: episode: 887, duration: 25.126s, episode steps: 2139, steps per second: 85, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2319614/3500000: episode: 888, duration: 23.525s, episode steps: 2002, steps per second: 85, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean action: 2321373/3500000: episode: 889, duration: 20.696s, episode steps: 1759, steps per second: 85, episode reward: 20.000, mean reward: 0.011 [-1.000, 1.000], mean action: 2323811/3500000: episode: 890, duration: 28.645s, episode steps: 2438, steps per second: 85, episode reward: 18.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2325899/3500000: episode: 891, duration: 24.629s, episode steps: 2088, steps per second: 85, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean action: 2328564/3500000: episode: 892, duration: 31.305s, episode steps: 2665, steps per second: 85, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2331044/3500000: episode: 893, duration: 29.632s, episode steps: 2480, steps per second: 84, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean action: 2333086/3500000: episode: 894, duration: 23.924s, episode steps: 2042, steps per second: 85, episode reward: 19.000, mean reward: 0.009 [-1.000, 1.000], mean action: 2335431/3500000: episode: 895, duration: 27.651s, episode steps: 2345, steps per second: 85, episode reward: 18.000, mean reward: 0.008 [-1.000, 1.000], mean action: 2337781/3500000: episode: 896, duration: 27.746s, episode steps: 2350, steps per second: 85, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2340831/3500000: episode: 897, duration: 36.045s, episode steps: 3050, steps per second: 85, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action: 2343362/3500000: episode: 898, duration: 30.528s, episode steps: 2531, steps per second: 83, episode reward: 17.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2346744/3500000: episode: 899, duration: 39.646s, episode steps: 3382, steps per second: 85, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action: 2349538/3500000: episode: 900, duration: 32.180s, episode steps: 2794, steps per second: 85, episode reward: 15.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2351461/3500000: episode: 901, duration: 54.248s, episode steps: 4623, steps per second: 85, episode reward: 6.000, mean reward: 0.001 [-1.000, 1.000], mean action: 2356731/3500000: episode: 902, duration: 29.913s, episode steps: 2570, steps per second: 86, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2360807/3500000: episode: 903, duration: 47.690s, episode steps: 4076, steps per second: 85, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action: 2363386/3500000: episode: 904, duration: 30.100s, episode steps: 2579, steps per second: 86, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean action: 2365921/3500000: episode: 905, duration: 29.641s, episode steps: 2535, steps per second: 86, episode reward: 18.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2369867/3500000: episode: 906, duration: 45.965s, episode steps: 3946, steps per second: 86, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action: 2372003/3500000: episode: 907, duration: 24.963s, episode steps: 2136, steps per second: 86, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean action: 2374298/3500000: episode: 908, duration: 26.833s, episode steps: 2295, steps per second: 86, episode reward: 18.000, mean reward: 0.008 [-1.000, 1.000], mean action: 2376922/3500000: episode: 909, duration: 30.691s, episode steps: 2624, steps per second: 85, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean action: 2380823/3500000: episode: 910, duration: 45.562s, episode steps: 3901, steps per second: 86, episode reward: 13.000, mean reward: 0.003 [-1.000, 1.000], mean action: 2382804/3500000: episode: 911, duration: 32.173s, episode steps: 1981, steps per second: 85, episode reward: 21.000, mean reward: 0.011 [0.000, 1.000], mean action: 2385561/3500000: episode: 912, duration: 32.180s, episode steps: 2757, steps per second: 86, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean action: 2388095/3500000: episode: 913, duration: 29.490s, episode steps: 2534, steps per second: 86, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean action: 2390528/3500000: episode: 914, duration: 28.443s, episode steps: 2433, steps per second: 86, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2393363/3500000: episode: 915, duration: 33.264s, episode steps: 2835, steps per second: 85, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2395717/3500000: episode: 916, duration: 27.555s, episode steps: 2354, steps per second: 85, episode reward: 17.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2398154/3500000: episode: 917, duration: 28.550s, episode steps: 2437, steps per second: 85, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2400547/3500000: episode: 918, duration: 28.049s, episode steps: 2393, steps per second: 85, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2402887/3500000: episode: 919, duration: 27.555s, episode steps: 2340, steps per second: 85, episode reward: 18.000, mean reward: 0.008 [-1.000, 1.000], mean action: 2405344/3500000: episode: 920, duration: 28.934s, episode steps: 2457, steps per second: 85, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean action: 2407896/3500000: episode: 921, duration: 30.277s, episode steps: 2552, steps per second: 84, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2410795/3500000: episode: 922, duration: 34.842s, episode steps: 2899, steps per second: 83, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2412528/3500000: episode: 923, duration: 20.521s, episode steps: 1733, steps per second: 84, episode reward: 19.000, mean reward: 0.011 [-1.000, 1.000], mean action: 2416205/3500000: episode: 924, duration: 42.954s, episode steps: 3677, steps per second: 86, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action: 2418747/3500000: episode: 925, duration: 29.659s, episode steps: 2542, steps per second: 86, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean action: 2421872/3500000: episode: 926, duration: 36.547s, episode steps: 3125, steps per second: 86, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action: 2424354/3500000: episode: 927, duration: 28.955s, episode steps: 2482, steps per second: 86, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2427818/3500000: episode: 928, duration: 40.408s, episode steps: 3464, steps per second: 86, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action: 2430039/3500000: episode: 929, duration: 25.907s, episode steps: 2221, steps per second: 86, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2432135/3500000: episode: 930, duration: 24.582s, episode steps: 2096, steps per second: 85, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean action: 2434128/3500000: episode: 931, duration: 23.352s, episode steps: 1993, steps per second: 85, episode reward: 17.000, mean reward: 0.009 [-1.000, 1.000], mean action: 2436253/3500000: episode: 932, duration: 24.920s, episode steps: 2125, steps per second: 85, episode reward: 16.000, mean reward: 0.008 [-1.000, 1.000], mean action: 2439245/3500000: episode: 933, duration: 34.969s, episode steps: 2992, steps per second: 86, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean action: 2442426/3500000: episode: 934, duration: 37.266s, episode steps: 3181, steps per second: 85, episode reward: 13.000, mean reward: 0.004 [-1.000, 1.000], mean action: 2445243/3500000: episode: 935, duration: 32.871s, episode steps: 2817, steps per second: 86, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean action: 2447983/3500000: episode: 936, duration: 32.019s, episode steps: 2740, steps per second: 86, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean action: 2451457/3500000: episode: 937, duration: 40.555s, episode steps: 3474, steps per second: 86, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action: 2453775/3500000: episode: 938, duration: 27.172s, episode steps: 2318, steps per second: 85, episode reward: 17.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2456128/3500000: episode: 939, duration: 27.403s, episode steps: 2353, steps per second: 86, episode reward: 17.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2458482/3500000: episode: 940, duration: 27.589s, episode steps: 2354, steps per second: 85, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2460802/3500000: episode: 941, duration: 27.092s, episode steps: 2320, steps per second: 86, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2463663/3500000: episode: 942, duration: 33.655s, episode steps: 2861, steps per second: 85, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean action: 2465779/3500000: episode: 943, duration: 24.619s, episode steps: 2116, steps per second: 86, episode reward: 16.000, mean reward: 0.008 [-1.000, 1.000], mean action: 2467920/3500000: episode: 944, duration: 25.056s, episode steps: 2141, steps per second: 85, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2470053/3500000: episode: 945, duration: 25.000s, episode steps: 2133, steps per second: 85, episode reward: 16.000, mean reward: 0.008 [-1.000, 1.000], mean action: 2472437/3500000: episode: 946, duration: 27.911s, episode steps: 2384, steps per second: 85, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2474956/3500000: episode: 947, duration: 29.526s, episode steps: 2519, steps per second: 85, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2478464/3500000: episode: 948, duration: 41.215s, episode steps: 3508, steps per second: 85, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action: 2481202/3500000: episode: 949, duration: 32.040s, episode steps: 2738, steps per second: 85, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean action: 2483821/3500000: episode: 950, duration: 30.877s, episode steps: 2619, steps per second: 85, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2486494/3500000: episode: 951, duration: 31.268s, episode steps: 2673, steps per second: 85, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean action: 2489007/3500000: episode: 952, duration: 29.417s, episode steps: 2513, steps per second: 85, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2490982/3500000: episode: 953, duration: 23.105s, episode steps: 1975, steps per second: 85, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean action: 2493769/3500000: episode: 954, duration: 32.661s, episode steps: 2787, steps per second: 85, episode reward: 7.000, mean reward: 0.003 [-1.000, 1.000], mean action: 2496140/3500000: episode: 955, duration: 27.727s, episode steps: 2371, steps per second: 86, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2498886/3500000: episode: 956, duration: 32.148s, episode steps: 2746, steps per second: 85, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2502255/3500000: episode: 957, duration: 39.340s, episode steps: 3369, steps per second: 86, episode reward: 11.000, mean reward: 0.003 [-1.000, 1.000], mean action: 2505005/3500000: episode: 958, duration: 32.430s, episode steps: 2750, steps per second: 85, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2506881/3500000: episode: 959, duration: 22.058s, episode steps: 1876, steps per second: 85, episode reward: 19.000, mean reward: 0.010 [-1.000, 1.000], mean action: 2509659/3500000: episode: 960, duration: 32.552s, episode steps: 2778, steps per second: 85, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action: 2512465/3500000: episode: 961, duration: 32.903s, episode steps: 2806, steps per second: 85, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean action: 2515359/3500000: episode: 962, duration: 33.912s, episode steps: 2894, steps per second: 85, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action: 2517697/3500000: episode: 963, duration: 27.288s, episode steps: 2338, steps per second: 86, episode reward: 17.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2520104/3500000: episode: 964, duration: 28.169s, episode steps: 2407, steps per second: 85, episode reward: 18.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2522022/3500000: episode: 965, duration: 22.504s, episode steps: 1918, steps per second: 85, episode reward: 19.000, mean reward: 0.010 [-1.000, 1.000], mean action: 2524696/3500000: episode: 966, duration: 31.314s, episode steps: 2674, steps per second: 85, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2527279/3500000: episode: 967, duration: 30.130s, episode steps: 2583, steps per second: 86, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2529982/3500000: episode: 968, duration: 31.591s, episode steps: 2703, steps per second: 86, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action: 2532519/3500000: episode: 969, duration: 29.633s, episode steps: 2537, steps per second: 86, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2535097/3500000: episode: 970, duration: 30.221s, episode steps: 2578, steps per second: 85, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2537196/3500000: episode: 971, duration: 24.536s, episode steps: 2099, steps per second: 86, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2539750/3500000: episode: 972, duration: 29.860s, episode steps: 2554, steps per second: 86, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action: 2542614/3500000: episode: 973, duration: 33.409s, episode steps: 2864, steps per second: 86, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean action: 2544982/3500000: episode: 974, duration: 27.817s, episode steps: 2368, steps per second: 85, episode reward: 18.000, mean reward: 0.008 [-1.000, 1.000], mean action: 2547183/3500000: episode: 975, duration: 25.708s, episode steps: 2201, steps per second: 86, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2549552/3500000: episode: 976, duration: 27.759s, episode steps: 2369, steps per second: 85, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean action: 2551558/3500000: episode: 977, duration: 23.383s, episode steps: 2006, steps per second: 86, episode reward: 21.000, mean reward: 0.010 [0.000, 1.000], mean action: 2553815/3500000: episode: 978, duration: 26.373s, episode steps: 2257, steps per second: 86, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2556097/3500000: episode: 979, duration: 26.831s, episode steps: 2282, steps per second: 85, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean action: 2558301/3500000: episode: 980, duration: 26.225s, episode steps: 2204, steps per second: 84, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action:

2560897/3500000:	episode:	981,	duration:	30.644s,	episode steps:	2596,	steps per second:	85,	episode reward:	11.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
2563219/3500000:	episode:	982,	duration:	27.227s,	episode steps:	2322,	steps per second:	85,	episode reward:	15.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
2565160/3500000:	episode:	983,	duration:	23.613s,	episode steps:	1941,	steps per second:	82,	episode reward:	19.000,	mean reward:	0.010	[-1.000, 1.000],	mean action:
2567412/3500000:	episode:	984,	duration:	26.670s,	episode steps:	2252,	steps per second:	84,	episode reward:	18.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
2570432/3500000:	episode:	985,	duration:	35.461s,	episode steps:	3020,	steps per second:	85,	episode reward:	8.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
2572727/3500000:	episode:	986,	duration:	26.910s,	episode steps:	2295,	steps per second:	85,	episode reward:	17.000,	mean reward:	0.007	[-1.000, 1.000],	mean action:
2574828/3500000:	episode:	987,	duration:	24.764s,	episode steps:	2101,	steps per second:	85,	episode reward:	18.000,	mean reward:	0.009	[-1.000, 1.000],	mean action:
2577102/3500000:	episode:	988,	duration:	26.961s,	episode steps:	2274,	steps per second:	84,	episode reward:	15.000,	mean reward:	0.007	[-1.000, 1.000],	mean action:
2580019/3500000:	episode:	989,	duration:	34.813s,	episode steps:	2917,	steps per second:	84,	episode reward:	11.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
2583168/3500000:	episode:	990,	duration:	36.892s,	episode steps:	3149,	steps per second:	85,	episode reward:	11.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
2585567/3500000:	episode:	991,	duration:	28.183s,	episode steps:	2399,	steps per second:	85,	episode reward:	17.000,	mean reward:	0.007	[-1.000, 1.000],	mean action:
2588036/3500000:	episode:	992,	duration:	28.983s,	episode steps:	2469,	steps per second:	85,	episode reward:	13.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
2590542/3500000:	episode:	993,	duration:	30.339s,	episode steps:	2506,	steps per second:	83,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
2593079/3500000:	episode:	994,	duration:	29.759s,	episode steps:	2537,	steps per second:	85,	episode reward:	15.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
2595178/3500000:	episode:	995,	duration:	24.679s,	episode steps:	2099,	steps per second:	85,	episode reward:	16.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
2597430/3500000:	episode:	996,	duration:	26.304s,	episode steps:	2252,	steps per second:	86,	episode reward:	17.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
2599710/3500000:	episode:	997,	duration:	26.654s,	episode steps:	2280,	steps per second:	86,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
2602549/3500000:	episode:	998,	duration:	33.184s,	episode steps:	2839,	steps per second:	86,	episode reward:	16.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
2605386/3500000:	episode:	999,	duration:	33.198s,	episode steps:	2837,	steps per second:	85,	episode reward:	14.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
2607702/3500000:	episode:	1000,	duration:	27.085s,	episode steps:	2316,	steps per second:	86,	episode reward:	16.000,	mean reward:	0.007	[-1.000, 1.000],	mean action:
2610714/3500000:	episode:	1001,	duration:	35.173s,	episode steps:	3012,	steps per second:	86,	episode reward:	10.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
2613688/3500000:	episode:	1002,	duration:	34.615s,	episode steps:	2974,	steps per second:	86,	episode reward:	9.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
2616073/3500000:	episode:	1003,	duration:	27.982s,	episode steps:	2385,	steps per second:	85,	episode reward:	15.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
2619027/3500000:	episode:	1004,	duration:	34.559s,	episode steps:	2954,	steps per second:	85,	episode reward:	13.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
2620796/3500000:	episode:	1005,	duration:	20.740s,	episode steps:	1769,	steps per second:	85,	episode reward:	21.000,	mean reward:	0.012	[0.000, 1.000],	mean action:
2623911/3500000:	episode:	1006,	duration:	37.775s,	episode steps:	3115,	steps per second:	82,	episode reward:	11.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
2626629/3500000:	episode:	1007,	duration:	32.192s,	episode steps:	2718,	steps per second:	84,	episode reward:	15.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
2629635/3500000:	episode:	1008,	duration:	35.399s,	episode steps:	3006,	steps per second:	85,	episode reward:	13.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
2632820/3500000:	episode:	1009,	duration:	37.295s,	episode steps:	3185,	steps per second:	85,	episode reward:	7.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
2634630/3500000:	episode:	1010,	duration:	21.215s,	episode steps:	1810,	steps per second:	85,	episode reward:	18.000,	mean reward:	0.010	[-1.000, 1.000],	mean action:
2637178/3500000:	episode:	1011,	duration:	30.292s,	episode steps:	2548,	steps per second:	84,	episode reward:	13.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
2639847/3500000:	episode:	1012,	duration:	31.101s,	episode steps:	2669,	steps per second:	86,	episode reward:	14.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
2642186/3500000:	episode:	1013,	duration:	27.445s,	episode steps:	2339,	steps per second:	85,	episode reward:	15.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
2644710/3500000:	episode:	1014,	duration:	29.561s,	episode steps:	2524,	steps per second:	85,	episode reward:	13.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
2647184/3500000:	episode:	1015,	duration:	29.144s,	episode steps:	2474,	steps per second:	85,	episode reward:	16.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
2649814/3500000:	episode:	1016,	duration:	30.893s,	episode steps:	2630,	steps per second:	85,	episode reward:	14.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
2652470/3500000:	episode:	1017,	duration:	31.121s,	episode steps:	2656,	steps per second:	85,	episode reward:	12.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
2654916/3500000:	episode:	1018,	duration:	28.606s,	episode steps:	2446,	steps per second:	86,	episode reward:	11.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
2656931/3500000:	episode:	1019,	duration:	23.657s,	episode steps:	2015,	steps per second:	85,	episode reward:	16.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
2658805/3500000:	episode:	1020,	duration:	21.860s,	episode steps:	1874,	steps per second:	86,	episode reward:	18.000,	mean reward:	0.010	[-1.000, 1.000],	mean action:
2661302/3500000:	episode:	1021,	duration:	29.194s,	episode steps:	2497,	steps per second:	86,	episode reward:	17.000,	mean reward:	0.007	[-1.000, 1.000],	mean action:
2663685/3500000:	episode:	1022,	duration:	27.942s,	episode steps:	2383,	steps per second:	85,	episode reward:	15.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
2665629/3500000:	episode:	1023,	duration:	22.708s,	episode steps:	1944,	steps per second:	86,	episode reward:	16.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
2667956/3500000:	episode:	1024,	duration:	27.281s,	episode steps:	2327,	steps per second:	85,	episode reward:	18.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
2670292/3500000:	episode:	1025,	duration:	27.449s,	episode steps:	2336,	steps per second:	85,	episode reward:	15.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
2672177/3500000:	episode:	1026,	duration:	22.055s,	episode steps:	1885,	steps per second:	85,	episode reward:	18.000,	mean reward:	0.010	[-1.000, 1.000],	mean action:
2674213/3500000:	episode:	1027,	duration:	23.815s,	episode steps:	2036,	steps per second:	85,	episode reward:	17.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
2677248/3500000:	episode:	1028,	duration:	35.578s,	episode steps:	3035,	steps per second:	85,	episode reward:	8.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
2680799/3500000:	episode:	1029,	duration:	41.425s,	episode steps:	3551,	steps per second:	86,	episode reward:	2.000,	mean reward:	0.001	[-1.000, 1.000],	mean action:
2682979/3500000:	episode:	1030,	duration:	25.523s,	episode steps:	2180,	steps per second:	85,	episode reward:	15.000,	mean reward:	0.007	[-1.000, 1.000],	mean action:
2685087/3500000:	episode:	1031,	duration:	24.648s,	episode steps:	2108,	steps per second:	86,	episode reward:	19.000,	mean reward:	0.009	[-1.000, 1.000],	mean action:
2687125/3500000:	episode:	1032,	duration:	23.965s,	episode steps:	2038,	steps per second:	85,	episode reward:	19.000,	mean reward:	0.009	[-1.000, 1.000],	mean action:
2689181/3500000:	episode:	1033,	duration:	24.024s,	episode steps:	2056,	steps per second:	86,	episode reward:	16.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
2691308/3500000:	episode:	1034,	duration:	24.895s,	episode steps:	2127,	steps per second:	85,	episode reward:	16.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
2693849/3500000:	episode:	1035,	duration:	29.734s,	episode steps:	2541,	steps per second:	85,	episode reward:	11.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
2695946/3500000:	episode:	1036,	duration:	24.451s,	episode steps:	2097,	steps per second:	86,	episode reward:	18.000,	mean reward:	0.009	[-1.000, 1.000],	mean action:
2698570/3500000:	episode:	1037,	duration:	42.496s,	episode steps:	3624,	steps per second:	85,	episode reward:	1.000,	mean reward:	0.000	[-1.000, 1.000],	mean action:
2701768/3500000:	episode:	1038,	duration:	25.750s,	episode steps:	2198,	steps per second:	85,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
2703790/3500000:	episode:	1039,	duration:	23.695s,	episode steps:	2022,	steps per second:	85,	episode reward:	17.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
2706649/3500000:	episode:	1040,	duration:	33.345s,	episode steps:	2859,	steps per second:	86,	episode reward:	8.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
2710355/3500000:	episode:	1041,	duration:	43.345s,	episode steps:	3706,	steps per second:	86,	episode reward:	7.000,	mean reward:	0.002	[-1.000, 1.000],	mean action:
2713173/3500000:	episode:	1042,	duration:	33.041s,	episode steps:	2818,	steps per second:	85,	episode reward:	15.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
2715499/3500000:	episode:	1043,	duration:	27.153s,	episode steps:	2326,	steps per second:	86,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
2717785/3500000:	episode:	1044,	duration:	26.864s,	episode steps:	2286,	steps per second:	85,	episode reward:	20.000,	mean reward:	0.009	[-1.000, 1.000],	mean action:
2720532/3500000:	episode:	1045,	duration:	32.223s,	episode steps:	2747,	steps per second:	85,	episode reward:	12.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
2723004/3500000:	episode:	1046,	duration:	28.872s,	episode steps:	2472,	steps per second:	86,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
2725613/3500000:	episode:	1047,	duration:	30.687s,	episode steps:	2609,	steps per second:	85,	episode reward:	15.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
2728941/3500000:	episode:	1048,	duration:	38.927s,	episode steps:	3328,	steps per second:	85,	episode reward:	10.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
2731127/3500000:	episode:	1049,	duration:	25.490s,	episode steps:	2186,	steps per second:	86,	episode reward:	17.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
2733742/3500000:	episode:	1050,	duration:	30.533s,	episode steps:	2615,	steps per second:	86,	episode reward:	15.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
2736203/3500000:	episode:	1051,	duration:	28.773s,	episode steps:	2461,	steps per second:	86,	episode reward:	12.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
2738331/3500000:	episode:	1052,	duration:	24.936s,	episode steps:	2128,	steps per second:	85,	episode reward:	16.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
2740278/3500000:	episode:	1053,	duration:	22.805s,	episode steps:	1947,	steps per second:	85,	episode reward:	19.000,	mean reward:	0.010	[-1.000, 1.000],	mean action:
2742957/3500000:	episode:	1054,	duration:	31.316s,	episode steps:	2679,	steps per second:	86,	episode reward:	11.000,	mean reward:	0.004	[-1.000, 1.000],	mean action:
2745382/3500000:	episode:	1055,	duration:	28.324s,	episode steps:	2425,	steps per second:	86,	episode reward:	13.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
2747277/3500000:	episode:	1056,	duration:	22.219s,	episode steps:	1895,	steps per second:	85,	episode reward:	17.000,	mean reward:	0.009	[-1.000, 1.000],	mean action:
2749564/3500000:	episode:	1057,	duration:	26.811s,	episode steps:	2287,	steps per second:	85,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
2751670/3500000:	episode:	1058,	duration:	24.515s,	episode steps:	2106,	steps per second:	86,	episode reward:	17.000,	mean reward:	0.008	[-1.000, 1.000],	mean action:
2754038/3500000:	episode:	1059,	duration:	27.723s,	episode steps:	2368,	steps per second:	85,	episode reward:	14.000,	mean reward:	0.006	[-1.000, 1.000],	mean action:
2757011/3500000:	episode:	1060,	duration:	34.734s,	episode steps:	2973,	steps per second:	86,	episode reward:	10.000,	mean reward:	0.003	[-1.000, 1.000],	mean action:
2759862/3500000:	episode:	1061,	duration:	33.391s,	episode steps:	2851,	steps per second:	85,	episode reward:	14.000,	mean reward:	0.005	[-1.000, 1.000],	mean action:
2761508/3500000:	episode:	1062,	duration:	19.289s,	episode steps:	1646,	steps per second:	85,	episode reward:	21.000,	mean reward:	0.013	[0.000, 1.000],	mean action:
2763634/350														

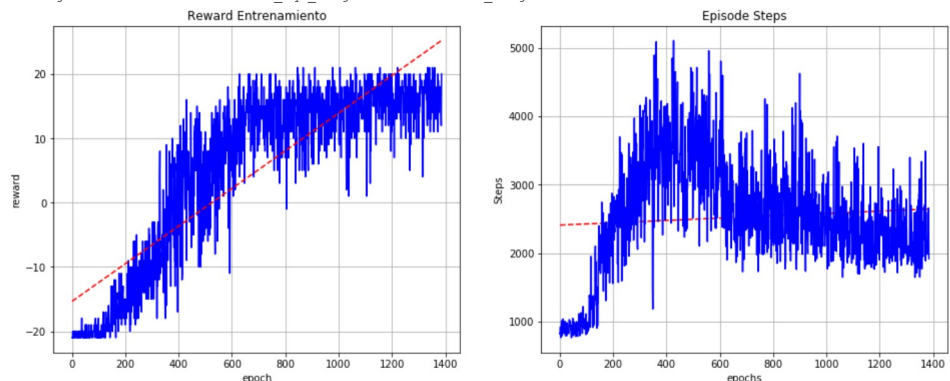

```

3325668/3500000: episode: 1308, duration: 22.181s, episode steps: 1875, steps per second: 85, episode reward: 19.000, mean reward: 0.010 [-1.000, 1.000], mean actor
3327774/3500000: episode: 1309, duration: 24.909s, episode steps: 2106, steps per second: 85, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean actor
3329678/3500000: episode: 1310, duration: 22.570s, episode steps: 1904, steps per second: 84, episode reward: 19.000, mean reward: 0.010 [-1.000, 1.000], mean actor
3331681/3500000: episode: 1311, duration: 23.726s, episode steps: 2003, steps per second: 84, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean actor
3333662/3500000: episode: 1312, duration: 23.441s, episode steps: 1981, steps per second: 85, episode reward: 17.000, mean reward: 0.009 [-1.000, 1.000], mean actor
3335620/3500000: episode: 1313, duration: 23.348s, episode steps: 1958, steps per second: 84, episode reward: 17.000, mean reward: 0.009 [-1.000, 1.000], mean actor
3339016/3500000: episode: 1314, duration: 40.114s, episode steps: 3396, steps per second: 85, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
3341480/3500000: episode: 1315, duration: 29.096s, episode steps: 2464, steps per second: 85, episode reward: 16.000, mean reward: 0.006 [-1.000, 1.000], mean actor
3343441/3500000: episode: 1316, duration: 23.116s, episode steps: 1961, steps per second: 85, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean actor
3345698/3500000: episode: 1317, duration: 26.546s, episode steps: 2257, steps per second: 85, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean actor
3348395/3500000: episode: 1318, duration: 31.854s, episode steps: 2697, steps per second: 85, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean actor
3350918/3500000: episode: 1319, duration: 29.804s, episode steps: 2523, steps per second: 85, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean actor
3353357/3500000: episode: 1320, duration: 28.684s, episode steps: 2439, steps per second: 85, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean actor
3355179/3500000: episode: 1321, duration: 21.378s, episode steps: 1822, steps per second: 85, episode reward: 19.000, mean reward: 0.010 [-1.000, 1.000], mean actor
3357492/3500000: episode: 1322, duration: 27.255s, episode steps: 2313, steps per second: 85, episode reward: 17.000, mean reward: 0.007 [-1.000, 1.000], mean actor
3359579/3500000: episode: 1323, duration: 24.935s, episode steps: 2087, steps per second: 84, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean actor
3361626/3500000: episode: 1324, duration: 24.084s, episode steps: 2047, steps per second: 85, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean actor
3363504/3500000: episode: 1325, duration: 22.098s, episode steps: 1878, steps per second: 85, episode reward: 20.000, mean reward: 0.011 [-1.000, 1.000], mean actor
3365385/3500000: episode: 1326, duration: 22.241s, episode steps: 1881, steps per second: 85, episode reward: 19.000, mean reward: 0.010 [-1.000, 1.000], mean actor
3367263/3500000: episode: 1327, duration: 22.134s, episode steps: 1878, steps per second: 85, episode reward: 20.000, mean reward: 0.011 [-1.000, 1.000], mean actor
3369238/3500000: episode: 1328, duration: 23.428s, episode steps: 1975, steps per second: 84, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean actor
3371402/3500000: episode: 1329, duration: 25.463s, episode steps: 2164, steps per second: 85, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean actor
3373241/3500000: episode: 1330, duration: 21.799s, episode steps: 1839, steps per second: 84, episode reward: 20.000, mean reward: 0.011 [-1.000, 1.000], mean actor
3375827/3500000: episode: 1331, duration: 30.387s, episode steps: 2586, steps per second: 85, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean actor
3378193/3500000: episode: 1332, duration: 27.958s, episode steps: 2366, steps per second: 85, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean actor
3379838/3500000: episode: 1333, duration: 19.321s, episode steps: 1645, steps per second: 85, episode reward: 21.000, mean reward: 0.013 [0.000, 1.000], mean action:
3381791/3500000: episode: 1334, duration: 22.874s, episode steps: 1953, steps per second: 85, episode reward: 19.000, mean reward: 0.010 [-1.000, 1.000], mean actor
3383838/3500000: episode: 1335, duration: 23.932s, episode steps: 2047, steps per second: 86, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean actor
3386137/3500000: episode: 1336, duration: 26.893s, episode steps: 2299, steps per second: 85, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean actor
3388739/3500000: episode: 1337, duration: 30.427s, episode steps: 2602, steps per second: 86, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean actor
3390531/3500000: episode: 1338, duration: 20.972s, episode steps: 1792, steps per second: 85, episode reward: 20.000, mean reward: 0.011 [-1.000, 1.000], mean actor
3392443/3500000: episode: 1339, duration: 22.414s, episode steps: 1912, steps per second: 85, episode reward: 19.000, mean reward: 0.010 [-1.000, 1.000], mean actor
3394152/3500000: episode: 1340, duration: 19.995s, episode steps: 1709, steps per second: 85, episode reward: 21.000, mean reward: 0.012 [0.000, 1.000], mean action:
3396384/3500000: episode: 1341, duration: 26.121s, episode steps: 2232, steps per second: 85, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean actor
3398659/3500000: episode: 1342, duration: 26.541s, episode steps: 2275, steps per second: 86, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean actor
3401345/3500000: episode: 1343, duration: 31.505s, episode steps: 2686, steps per second: 85, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean actor
3403502/3500000: episode: 1344, duration: 25.206s, episode steps: 2157, steps per second: 86, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean actor
3405369/3500000: episode: 1345, duration: 21.785s, episode steps: 1867, steps per second: 86, episode reward: 18.000, mean reward: 0.010 [-1.000, 1.000], mean actor
3408140/3500000: episode: 1346, duration: 32.363s, episode steps: 2771, steps per second: 86, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean actor
3409916/3500000: episode: 1347, duration: 20.868s, episode steps: 1776, steps per second: 85, episode reward: 19.000, mean reward: 0.011 [-1.000, 1.000], mean actor
3412780/3500000: episode: 1348, duration: 33.585s, episode steps: 2864, steps per second: 85, episode reward: 16.000, mean reward: 0.006 [-1.000, 1.000], mean actor
3415159/3500000: episode: 1349, duration: 27.779s, episode steps: 2379, steps per second: 86, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean actor
3418242/3500000: episode: 1350, duration: 36.073s, episode steps: 3083, steps per second: 85, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean actor
3419892/3500000: episode: 1351, duration: 19.351s, episode steps: 1650, steps per second: 85, episode reward: 21.000, mean reward: 0.013 [0.000, 1.000], mean action:
3421691/3500000: episode: 1352, duration: 20.994s, episode steps: 1799, steps per second: 86, episode reward: 20.000, mean reward: 0.011 [-1.000, 1.000], mean actor
3423890/3500000: episode: 1353, duration: 25.757s, episode steps: 2199, steps per second: 85, episode reward: 20.000, mean reward: 0.009 [-1.000, 1.000], mean actor
3425791/3500000: episode: 1354, duration: 22.321s, episode steps: 1901, steps per second: 85, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean actor
3428166/3500000: episode: 1355, duration: 27.903s, episode steps: 2375, steps per second: 85, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean actor
3430222/3500000: episode: 1356, duration: 24.144s, episode steps: 2056, steps per second: 85, episode reward: 19.000, mean reward: 0.009 [-1.000, 1.000], mean actor
3433557/3500000: episode: 1357, duration: 38.952s, episode steps: 3335, steps per second: 86, episode reward: 11.000, mean reward: 0.003 [-1.000, 1.000], mean actor
3436068/3500000: episode: 1358, duration: 29.349s, episode steps: 2511, steps per second: 86, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean actor
3437837/3500000: episode: 1359, duration: 20.798s, episode steps: 1769, steps per second: 85, episode reward: 21.000, mean reward: 0.012 [0.000, 1.000], mean action:
3440298/3500000: episode: 1360, duration: 29.001s, episode steps: 2461, steps per second: 85, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean actor
3442653/3500000: episode: 1361, duration: 27.595s, episode steps: 2355, steps per second: 85, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean actor
3444895/3500000: episode: 1362, duration: 26.321s, episode steps: 2242, steps per second: 85, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean actor
3446847/3500000: episode: 1363, duration: 22.982s, episode steps: 1952, steps per second: 85, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean actor
3449241/3500000: episode: 1364, duration: 28.068s, episode steps: 2394, steps per second: 85, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean actor
3451445/3500000: episode: 1365, duration: 25.941s, episode steps: 2204, steps per second: 85, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean actor
3454112/3500000: episode: 1366, duration: 31.266s, episode steps: 2667, steps per second: 85, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean actor
3456739/3500000: episode: 1367, duration: 30.842s, episode steps: 2627, steps per second: 85, episode reward: 16.000, mean reward: 0.006 [-1.000, 1.000], mean actor
3459288/3500000: episode: 1368, duration: 29.925s, episode steps: 2549, steps per second: 85, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean actor
3461484/3500000: episode: 1369, duration: 25.884s, episode steps: 2196, steps per second: 85, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean actor
3464193/3500000: episode: 1370, duration: 31.793s, episode steps: 2709, steps per second: 85, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean actor
3466099/3500000: episode: 1371, duration: 22.352s, episode steps: 1906, steps per second: 85, episode reward: 20.000, mean reward: 0.010 [-1.000, 1.000], mean actor
3469587/3500000: episode: 1372, duration: 40.844s, episode steps: 3488, steps per second: 85, episode reward: 11.000, mean reward: 0.003 [-1.000, 1.000], mean actor
3471474/3500000: episode: 1373, duration: 22.233s, episode steps: 1887, steps per second: 85, episode reward: 19.000, mean reward: 0.010 [-1.000, 1.000], mean actor
3474074/3500000: episode: 1374, duration: 30.442s, episode steps: 2600, steps per second: 85, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean actor
3476655/3500000: episode: 1375, duration: 30.245s, episode steps: 2581, steps per second: 85, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean actor
3478840/3500000: episode: 1376, duration: 25.657s, episode steps: 2185, steps per second: 85, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean actor
3480887/3500000: episode: 1377, duration: 24.128s, episode steps: 2047, steps per second: 85, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean actor
3482883/3500000: episode: 1378, duration: 23.351s, episode steps: 1996, steps per second: 85, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean actor
3484872/3500000: episode: 1379, duration: 23.504s, episode steps: 1989, steps per second: 85, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean actor
3486854/3500000: episode: 1380, duration: 23.260s, episode steps: 1982, steps per second: 85, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean actor
3488913/3500000: episode: 1381, duration: 24.104s, episode steps: 2059, steps per second: 85, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean actor
3491084/3500000: episode: 1382, duration: 25.695s, episode steps: 2171, steps per second: 84, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean actor
3493470/3500000: episode: 1383, duration: 28.035s, episode steps: 2386, steps per second: 85, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean actor
3496125/3500000: episode: 1384, duration: 31.123s, episode steps: 2655, steps per second: 85, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean actor
3498041/3500000: episode: 1385, duration: 22.452s, episode steps: 1916, steps per second: 85, episode reward: 20.000, mean reward: 0.010 [-1.000, 1.000], mean actor

```

done, took 45348.288 seconds

Pesos guardados en: 202001170743_dqn_PongDeterministic-v0_weights.h5f



Testing for 10 episodes ...

```

Episode 1: reward: 16.000, steps: 2026
Episode 2: reward: 18.000, steps: 2109
Episode 3: reward: 20.000, steps: 1776
Episode 4: reward: 18.000, steps: 1988
Episode 5: reward: 18.000, steps: 2155
Episode 6: reward: 14.000, steps: 2842
Episode 7: reward: 17.000, steps: 1878

```

```
Episode 8: reward: 19.000, steps: 2074
Episode 9: reward: 13.000, steps: 2508
Episode 10: reward: 14.000, steps: 2710
mean_reward: 16.7
mean_steps: 2206.6
(16.7, 2206.6)
```