

```
Training for 1500000 steps ...
```

972/1500000:	episode: 1,	duration: 7.715s,	episode steps: 972,	steps per second: 125,	episode reward: -21.000,	mean reward: -0.022 [-1.000, 0.000],	mean action:
1886/1500000:	episode: 2,	duration: 2.112s,	episode steps: 914,	steps per second: 433,	episode reward: -21.000,	mean reward: -0.023 [-1.000, 0.000],	mean action:
3150/1500000:	episode: 3,	duration: 2.896s,	episode steps: 1264,	steps per second: 437,	episode reward: -19.000,	mean reward: -0.015 [-1.000, 1.000],	mean action:
3914/1500000:	episode: 4,	duration: 1.766s,	episode steps: 764,	steps per second: 433,	episode reward: -21.000,	mean reward: -0.027 [-1.000, 0.000],	mean action:
4905/1500000:	episode: 5,	duration: 2.247s,	episode steps: 991,	steps per second: 441,	episode reward: -20.000,	mean reward: -0.020 [-1.000, 1.000],	mean action:
5729/1500000:	episode: 6,	duration: 1.838s,	episode steps: 824,	steps per second: 448,	episode reward: -21.000,	mean reward: -0.025 [-1.000, 0.000],	mean action:
6710/1500000:	episode: 7,	duration: 2.253s,	episode steps: 981,	steps per second: 435,	episode reward: -20.000,	mean reward: -0.020 [-1.000, 1.000],	mean action:
7751/1500000:	episode: 8,	duration: 2.353s,	episode steps: 1041,	steps per second: 442,	episode reward: -20.000,	mean reward: -0.019 [-1.000, 1.000],	mean action:
8726/1500000:	episode: 9,	duration: 2.203s,	episode steps: 975,	steps per second: 442,	episode reward: -20.000,	mean reward: -0.021 [-1.000, 1.000],	mean action:
9550/1500000:	episode: 10,	duration: 1.840s,	episode steps: 824,	steps per second: 448,	episode reward: -21.000,	mean reward: -0.025 [-1.000, 0.000],	mean action:
10403/1500000:	episode: 11,	duration: 4.085s,	episode steps: 853,	steps per second: 209,	episode reward: -21.000,	mean reward: -0.025 [-1.000, 0.000],	mean action:
11227/1500000:	episode: 12,	duration: 4.544s,	episode steps: 824,	steps per second: 181,	episode reward: -21.000,	mean reward: -0.025 [-1.000, 0.000],	mean action:
12111/1500000:	episode: 13,	duration: 4.986s,	episode steps: 884,	steps per second: 177,	episode reward: -21.000,	mean reward: -0.024 [-1.000, 0.000],	mean action:
13197/1500000:	episode: 14,	duration: 5.975s,	episode steps: 1086,	steps per second: 182,	episode reward: -20.000,	mean reward: -0.018 [-1.000, 1.000],	mean action:
13961/1500000:	episode: 15,	duration: 4.174s,	episode steps: 764,	steps per second: 183,	episode reward: -21.000,	mean reward: -0.027 [-1.000, 0.000],	mean action:
14907/1500000:	episode: 16,	duration: 5.111s,	episode steps: 946,	steps per second: 185,	episode reward: -21.000,	mean reward: -0.022 [-1.000, 0.000],	mean action:
15699/1500000:	episode: 17,	duration: 4.287s,	episode steps: 792,	steps per second: 185,	episode reward: -21.000,	mean reward: -0.027 [-1.000, 0.000],	mean action:
16542/1500000:	episode: 18,	duration: 4.547s,	episode steps: 843,	steps per second: 185,	episode reward: -20.000,	mean reward: -0.024 [-1.000, 1.000],	mean action:
17486/1500000:	episode: 19,	duration: 5.186s,	episode steps: 944,	steps per second: 182,	episode reward: -19.000,	mean reward: -0.020 [-1.000, 1.000],	mean action:
18476/1500000:	episode: 20,	duration: 5.333s,	episode steps: 990,	steps per second: 186,	episode reward: -20.000,	mean reward: -0.020 [-1.000, 1.000],	mean action:
19300/1500000:	episode: 21,	duration: 4.531s,	episode steps: 824,	steps per second: 182,	episode reward: -21.000,	mean reward: -0.025 [-1.000, 0.000],	mean action:
20064/1500000:	episode: 22,	duration: 4.144s,	episode steps: 764,	steps per second: 184,	episode reward: -21.000,	mean reward: -0.027 [-1.000, 0.000],	mean action:
20884/1500000:	episode: 23,	duration: 4.577s,	episode steps: 820,	steps per second: 179,	episode reward: -21.000,	mean reward: -0.026 [-1.000, 0.000],	mean action:
21706/1500000:	episode: 24,	duration: 4.498s,	episode steps: 822,	steps per second: 183,	episode reward: -21.000,	mean reward: -0.026 [-1.000, 0.000],	mean action:
22532/1500000:	episode: 25,	duration: 4.496s,	episode steps: 826,	steps per second: 184,	episode reward: -21.000,	mean reward: -0.025 [-1.000, 0.000],	mean action:
23375/1500000:	episode: 26,	duration: 4.608s,	episode steps: 843,	steps per second: 183,	episode reward: -21.000,	mean reward: -0.025 [-1.000, 0.000],	mean action:
24328/1500000:	episode: 27,	duration: 5.240s,	episode steps: 953,	steps per second: 182,	episode reward: -20.000,	mean reward: -0.021 [-1.000, 1.000],	mean action:
25152/1500000:	episode: 28,	duration: 4.639s,	episode steps: 824,	steps per second: 178,	episode reward: -21.000,	mean reward: -0.025 [-1.000, 0.000],	mean action:
25916/1500000:	episode: 29,	duration: 4.281s,	episode steps: 764,	steps per second: 178,	episode reward: -21.000,	mean reward: -0.027 [-1.000, 0.000],	mean action:
26759/1500000:	episode: 30,	duration: 4.754s,	episode steps: 843,	steps per second: 177,	episode reward: -21.000,	mean reward: -0.025 [-1.000, 0.000],	mean action:
27707/1500000:	episode: 31,	duration: 5.197s,	episode steps: 948,	steps per second: 182,	episode reward: -19.000,	mean reward: -0.020 [-1.000, 1.000],	mean action:
28606/1500000:	episode: 32,	duration: 4.981s,	episode steps: 899,	steps per second: 180,	episode reward: -20.000,	mean reward: -0.022 [-1.000, 1.000],	mean action:
29398/1500000:	episode: 33,	duration: 4.560s,	episode steps: 792,	steps per second: 174,	episode reward: -21.000,	mean reward: -0.027 [-1.000, 0.000],	mean action:
30317/1500000:	episode: 34,	duration: 5.337s,	episode steps: 919,	steps per second: 172,	episode reward: -20.000,	mean reward: -0.022 [-1.000, 1.000],	mean action:

221888/1500000:	episode: 218,	duration: 8.020s,	episode steps: 1521,	steps per second: 190,	episode reward: -18.000,	mean reward: -0.012	[-1.000, 1.000],	mean acti
221889/1500000:	episode: 219,	duration: 6.929s,	episode steps: 1233,	steps per second: 190,	episode reward: -18.000,	mean reward: -0.015	[-1.000, 1.000],	mean acti
223478/1500000:	episode: 220,	duration: 8.386s,	episode steps: 1590,	steps per second: 190,	episode reward: -21.000,	mean reward: -0.013	[-1.000, 0.000],	mean acti
225114/1500000:	episode: 221,	duration: 8.492s,	episode steps: 1636,	steps per second: 193,	episode reward: -19.000,	mean reward: -0.012	[-1.000, 1.000],	mean acti
226892/1500000:	episode: 222,	duration: 9.293s,	episode steps: 1778,	steps per second: 191,	episode reward: -19.000,	mean reward: -0.011	[-1.000, 1.000],	mean acti
228754/1500000:	episode: 223,	duration: 9.698s,	episode steps: 1862,	steps per second: 192,	episode reward: -15.000,	mean reward: -0.008	[-1.000, 1.000],	mean acti
230296/1500000:	episode: 224,	duration: 8.051s,	episode steps: 1542,	steps per second: 192,	episode reward: -18.000,	mean reward: -0.012	[-1.000, 1.000],	mean acti
231518/1500000:	episode: 225,	duration: 6.656s,	episode steps: 1222,	steps per second: 192,	episode reward: -19.000,	mean reward: -0.016	[-1.000, 1.000],	mean acti
232410/1500000:	episode: 226,	duration: 6.770s,	episode steps: 1300,	steps per second: 192,	episode reward: -19.000,	mean reward: -0.015	[-1.000, 1.000],	mean acti
234079/1500000:	episode: 227,	duration: 6.583s,	episode steps: 1261,	steps per second: 192,	episode reward: -20.000,	mean reward: -0.016	[-1.000, 1.000],	mean acti
235650/1500000:	episode: 228,	duration: 8.203s,	episode steps: 1571,	steps per second: 192,	episode reward: -19.000,	mean reward: -0.012	[-1.000, 1.000],	mean acti
237088/1500000:	episode: 229,	duration: 7.528s,	episode steps: 1438,	steps per second: 191,	episode reward: -19.000,	mean reward: -0.013	[-1.000, 1.000],	mean acti
238160/1500000:	episode: 230,	duration: 5.676s,	episode steps: 1072,	steps per second: 189,	episode reward: -21.000,	mean reward: -0.020	[-1.000, 0.000],	mean acti
239682/1500000:	episode: 231,	duration: 7.913s,	episode steps: 1522,	steps per second: 192,	episode reward: -19.000,	mean reward: -0.012	[-1.000, 1.000],	mean acti
241236/1500000:	episode: 232,	duration: 8.089s,	episode steps: 1554,	steps per second: 192,	episode reward: -17.000,	mean reward: -0.011	[-1.000, 1.000],	mean acti
242733/1500000:	episode: 233,	duration: 7.649s,	episode steps: 1467,	steps per second: 192,	episode reward: -18.000,	mean reward: -0.012	[-1.000, 1.000],	mean acti
244203/1500000:	episode: 234,	duration: 7.995s,	episode steps: 1529,	steps per second: 191,	episode reward: -20.000,	mean reward: -0.013	[-1.000, 1.000],	mean acti
246288/1500000:	episode: 235,	duration: 10.687s,	episode steps: 2056,	steps per second: 192,	episode reward: -17.000,	mean reward: -0.008	[-1.000, 1.000],	mean acti
248340/1500000:	episode: 236,	duration: 10.879s,	episode steps: 2060,	steps per second: 189,	episode reward: -15.000,	mean reward: -0.007	[-1.000, 1.000],	mean acti
249420/1500000:	episode: 237,	duration: 5.705s,	episode steps: 1072,	steps per second: 188,	episode reward: -20.000,	mean reward: -0.019	[-1.000, 1.000],	mean acti
250903/1500000:	episode: 238,	duration: 7.796s,	episode steps: 1483,	steps per second: 190,	episode reward: -19.000,	mean reward: -0.013	[-1.000, 1.000],	mean acti
252798/1500000:	episode: 239,	duration: 10.103s,	episode steps: 1895,	steps per second: 188,	episode reward: -17.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
254455/1500000:	episode: 240,	duration: 8.965s,	episode steps: 1657,	steps per second: 188,	episode reward: -20.000,	mean reward: -0.012	[-1.000, 1.000],	mean acti
256319/1500000:	episode: 241,	duration: 9.805s,	episode steps: 1864,	steps per second: 190,	episode reward: -19.000,	mean reward: -0.010	[-1.000, 1.000],	mean acti
258444/1500000:	episode: 242,	duration: 11.143s,	episode steps: 2125,	steps per second: 191,	episode reward: -14.000,	mean reward: -0.007	[-1.000, 1.000],	mean acti
260029/1500000:	episode: 243,	duration: 9.283s,	episode steps: 1775,	steps per second: 191,	episode reward: -16.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
262202/1500000:	episode: 244,	duration: 10.421s,	episode steps: 1983,	steps per second: 190,	episode reward: -18.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
263779/1500000:	episode: 245,	duration: 8.578s,	episode steps: 1577,	steps per second: 184,	episode reward: -19.000,	mean reward: -0.012	[-1.000, 1.000],	mean acti
265478/1500000:	episode: 246,	duration: 8.999s,	episode steps: 1699,	steps per second: 189,	episode reward: -17.000,	mean reward: -0.010	[-1.000, 1.000],	mean acti
267321/1500000:	episode: 247,	duration: 9.207s,	episode steps: 1753,	steps per second: 190,	episode reward: -18.000,	mean reward: -0.010	[-1.000, 1.000],	mean acti
268782/1500000:	episode: 248,	duration: 8.164s,	episode steps: 1551,	steps per second: 190,	episode reward: -17.000,	mean reward: -0.011	[-1.000, 1.000],	mean acti
269997/1500000:	episode: 249,	duration: 6.348s,	episode steps: 1215,	steps per second: 191,	episode reward: -19.000,	mean reward: -0.016	[-1.000, 1.000],	mean acti
271539/1500000:	episode: 250,	duration: 8.100s,	episode steps: 1542,	steps per second: 190,	episode reward: -17.000,	mean reward: -0.011	[-1.000, 1.000],	mean acti
273200/1500000:	episode: 251,	duration: 8.298s,	episode steps: 1661,	steps per second: 190,	episode reward: -18.000,	mean reward: -0.011	[-1.000, 1.000],	mean acti
275045/1500000:	episode: 252,	duration: 9.665s,	episode steps: 1845,	steps per second: 191,	episode reward: -17.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
276571/1500000:	episode: 253,	duration: 8.090s,	episode steps: 1526,	steps per second: 189,	episode reward: -18.000,	mean reward: -0.012	[-1.000, 1.000],	mean acti
278285/1500000:	episode: 254,	duration: 8.991s,	episode steps: 1714,	steps per second: 191,	episode reward: -16.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
279881/1500000:	episode: 255,	duration: 8.842s,	episode steps: 1696,	steps per second: 192,	episode reward: -16.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
281004/1500000:	episode: 256,	duration: 5.420s,	episode steps: 1023,	steps per second: 189,	episode reward: -20.000,	mean reward: -0.020	[-1.000, 1.000],	mean acti
282798/1500000:	episode: 257,	duration: 9.417s,	episode steps: 1794,	steps per second: 191,	episode reward: -18.000,	mean reward: -0.010	[-1.000, 1.000],	mean acti
284525/1500000:	episode: 258,	duration: 9.249s,	episode steps: 1727,	steps per second: 187,	episode reward: -16.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
286472/1500000:	episode: 259,	duration: 10.170s,	episode steps: 1947,	steps per second: 191,	episode reward: -14.000,	mean reward: -0.007	[-1.000, 1.000],	mean acti
288031/1500000:	episode: 260,	duration: 8.175s,	episode steps: 1559,	steps per second: 191,	episode reward: -20.000,	mean reward: -0.013	[-1.000, 1.000],	mean acti
289881/1500000:	episode: 261,	duration: 9.668s,	episode steps: 1850,	steps per second: 191,	episode reward: -15.000,	mean reward: -0.008	[-1.000, 1.000],	mean acti
291115/1500000:	episode: 262,	duration: 6.490s,	episode steps: 1234,	steps per second: 190,	episode reward: -17.000,	mean reward: -0.014	[-1.000, 1.000],	mean acti
293049/1500000:	episode: 263,	duration: 10.179s,	episode steps: 1934,	steps per second: 190,	episode reward: -16.000,	mean reward: -0.008	[-1.000, 1.000],	mean acti
295398/1500000:	episode: 264,	duration: 12.351s,	episode steps: 2349,	steps per second: 190,	episode reward: -14.000,	mean reward: -0.006	[-1.000, 1.000],	mean acti
297763/1500000:	episode: 265,	duration: 12.454s,	episode steps: 2365,	steps per second: 190,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
299682/1500000:	episode: 266,	duration: 10.146s,	episode steps: 1919,	steps per second: 189,	episode reward: -17.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
300965/1500000:	episode: 267,	duration: 6.720s,	episode steps: 1283,	steps per second: 191,	episode reward: -18.000,	mean reward: -0.014	[-1.000, 1.000],	mean acti
302316/1500000:	episode: 268,	duration: 7.041s,	episode steps: 1351,	steps per second: 192,	episode reward: -20.000,	mean reward: -0.015	[-1.000, 1.000],	mean acti
303743/1500000:	episode: 269,	duration: 7.494s,	episode steps: 1427,	steps per second: 190,	episode reward: -17.000,	mean reward: -0.012	[-1.000, 1.000],	mean acti
305302/1500000:	episode: 270,	duration: 8.169s,	episode steps: 1559,	steps per second: 191,	episode reward: -18.000,	mean reward: -0.012	[-1.000, 1.000],	mean acti
307112/1500000:	episode: 271,	duration: 9.485s,	episode steps: 1810,	steps per second: 191,	episode reward: -17.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
308591/1500000:	episode: 272,	duration: 7.878s,	episode steps: 1479,	steps per second: 190,	episode reward: -20.000,	mean reward: -0.014	[-1.000, 1.000],	mean acti
310159/1500000:	episode: 273,	duration: 8.259s,	episode steps: 1568,	steps per second: 190,	episode reward: -20.000,	mean reward: -0.013	[-1.000, 1.000],	mean acti
311738/1500000:	episode: 274,	duration: 8.286s,	episode steps: 1579,	steps per second: 191,	episode reward: -15.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
313036/1500000:	episode: 275,	duration: 8.271s,	episode steps: 1568,	steps per second: 190,	episode reward: -20.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
314992/1500000:	episode: 276,	duration: 8.907s,	episode steps: 1686,	steps per second: 189,	episode reward: -18.000,	mean reward: -0.011	[-1.000, 1.000],	mean acti
316808/1500000:	episode: 277,	duration: 9.594s,	episode steps: 1816,	steps per second: 189,	episode reward: -17.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
318979/1500000:	episode: 278,	duration: 11.380s,	episode steps: 2171,	steps per second: 191,	episode reward: -15.000,	mean reward: -0.007	[-1.000, 1.000],	mean acti
320904/1500000:	episode: 279,	duration: 10.123s,	episode steps: 1925,	steps per second: 190,	episode reward: -16.000,	mean reward: -0.008	[-1.000, 1.000],	mean acti
322476/1500000:	episode: 280,	duration: 8.562s,	episode steps: 1572,	steps per second: 184,	episode reward: -20.000,	mean reward: -0.013	[-1.000, 1.000],	mean acti
325131/1500000:	episode: 281,	duration: 13.983s,	episode steps: 2655,	steps per second: 190,	episode reward: -11.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
326855/1500000:	episode: 282,	duration: 9.074s,	episode steps: 1724,	steps per second: 190,	episode reward: -17.000,	mean reward: -0.010	[-1.000, 1.000],	mean acti
328448/1500000:	episode: 283,	duration: 8.444s,	episode steps: 1593,	steps per second: 189,	episode reward: -17.000,	mean reward: -0.011	[-1.000, 1.000],	mean acti
330252/1500000:	episode: 284,	duration: 9.518s,	episode steps: 1804,	steps per second: 190,	episode reward: -18.000,	mean reward: -0.010	[-1.000, 1.000],	mean acti
332578/1500000:	episode: 285,	duration: 12.221s,	episode steps: 2326,	steps per second: 190,	episode reward: -14.000,	mean reward: -0.006	[-1.000, 1.000],	mean acti
334350/1500000:	episode: 286,	duration: 9.381s,	episode steps: 1772,	steps per second: 189,	episode reward: -18.000,	mean reward: -0.010	[-1.000, 1.000],	mean acti
336196/1500000:	episode: 287,	duration: 9.702s,	episode steps: 1846,	steps per second: 190,	episode reward: -16.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
338012/1500000:	episode: 288,	duration: 9.519s,	episode steps: 1816,	steps per second: 191,	episode reward: -19.000,	mean reward: -0.010	[-1.000, 1.000],	mean acti
339925/1500000:	episode: 289,	duration: 10.038s,	episode steps: 1913,	steps per second: 191,	episode reward: -17.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
342219/1500000:	episode: 290,	duration: 11.935s,	episode steps: 2294,	steps per second: 192,	episode reward: -16.000,	mean reward: -0.007	[-1.000, 1.000],	mean acti
344237/1500000:	episode: 291,	duration: 10.683s,	episode steps: 2018,	steps per second: 189,	episode reward: -16.000,	mean reward: -0.008	[-1.000, 1.000],	mean acti
345906/1500000:	episode: 292,	duration: 8.761s,	episode steps: 1669,	steps per second: 191,	episode reward: -18.000,	mean reward: -0.011	[-1.000, 1.000],	mean acti
347986/1500000:	episode: 293,	duration: 10.813s,	episode steps: 2080,	steps per second: 192,	episode reward: -18.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
350384/1500000:	episode: 294,	duration: 12.542s,	episode steps: 2398,	steps per second: 191,	episode reward: -12.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
352261/1500000:	episode: 295,	duration: 9.819s,	episode steps: 1877,	steps per second: 191,	episode reward: -18.000,	mean reward: -0.010	[-1.000, 1.000],	mean acti
354040/1500000:	episode: 296,	duration: 9.276s,	episode steps: 1779,	steps per second: 192,	episode reward: -18.000,	mean reward: -0.010	[-1.000, 1.000],	mean acti
356182/1500000:	episode: 297,	duration: 11.160s,	episode steps: 2142,	steps per second: 192,	episode reward: -17.000,	mean reward: -0.008	[-1.000, 1.000],	mean acti
358713/1500000:	episode: 298,	duration: 13.337s,	episode steps: 2531,	steps per second: 190,	episode reward: -12.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
360963/1500000:	episode: 299,	duration: 11.808s,	episode steps: 2250,	steps per second: 191,	episode reward: -15.000,	mean reward: -0.007	[-1.000, 1.000],	mean acti
362983/1500000:	episode: 300,	duration: 10.572s,	episode steps: 2020,	steps per second: 191,	episode reward: -17.000,	mean reward: -0.008	[-1.000, 1.000],	mean acti
364942/1500000:	episode: 301,	duration: 10.227s,	episode steps: 1959,	steps per second: 192,	episode reward: -16.000,	mean reward: -0.008	[-1.000, 1.000],	mean acti
366454/1500000:	episode: 302,	duration: 7.920s,	episode steps: 1512,	steps per second: 191,	episode reward: -19.000,	mean reward: -0.013	[-1.000, 1.000],	mean acti
368518/1500000:	episode: 303,	duration: 10.772s,	episode steps: 2064,	steps per second: 192,	episode reward: -14.000,	mean reward: -0.007	[-1.000, 1.000],	mean acti
370319/1500000:	episode: 304,	duration: 9.399s,	episode steps: 1801,	steps per second: 192,	episode reward: -17.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
372621/1500000:	episode: 305,	duration: 12.034s,	episode steps: 2302,	steps per second: 191,	episode reward: -12.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
374796/1500000:	episode: 306,	duration: 11.491s,	episode steps: 2175,	steps per second: 189,	episode reward: -17.000,	mean reward: -0.008	[-1.000, 1.000],	mean acti
376823/1500000:	episode: 307,	duration: 10.635s,	episode steps: 2027,	steps per second: 191,	episode reward: -19.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
378423/1500000:	episode: 308,	duration: 8.378s,	episode steps: 1600,	steps per second: 191,	episode reward: -19.000,	mean reward: -0.012	[-1.000, 1.000],	mean acti
380803/1500000:	episode: 309,	duration: 12.607s,	episode steps: 2380,	steps per second: 189,	episode reward: -15.000,	mean reward: -0.006	[-1.000, 1.000],	mean acti
382895/1500000:	episode: 310,	duration: 11.201s,	episode steps: 2092,	steps per second: 187,	episode reward: -17.000,	mean reward: -0.008	[-1.000, 1.000],	mean acti
384688/1500000:	episode: 311,	duration: 9.359s,	episode steps: 1793,	steps per second: 192,	episode reward: -17.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
386909/1500000:	episode: 312,	duration: 11.657s,	episode steps: 2221,	steps per second: 191,	episode reward: -16.000,	mean reward: -0.007	[-1.000, 1.000],	mean acti

118060/1500000:	episode: 327,	duration: 9.551s,	episode steps: 1440,	steps per second: 191,	episode reward: -20.000,	mean reward: -0.014 [-1.000, 1.000],	mean act
118885/1500000:	episode: 328,	duration: 9.585s,	episode steps: 1825,	steps per second: 190,	episode reward: -16.000,	mean reward: -0.009 [-1.000, 1.000],	mean act
120571/1500000:	episode: 329,	duration: 8.914s,	episode steps: 1686,	steps per second: 189,	episode reward: -18.000,	mean reward: -0.011 [-1.000, 1.000],	mean act
122822/1500000:	episode: 330,	duration: 11.826s,	episode steps: 2251,	steps per second: 190,	episode reward: -17.000,	mean reward: -0.008 [-1.000, 1.000],	mean act
124893/1500000:	episode: 331,	duration: 10.818s,	episode steps: 2071,	steps per second: 191,	episode reward: -14.000,	mean reward: -0.007 [-1.000, 1.000],	mean act
126347/1500000:	episode: 332,	duration: 7.605s,	episode steps: 1454,	steps per second: 191,	episode reward: -20.000,	mean reward: -0.014 [-1.000, 1.000],	mean act
128205/1500000:	episode: 333,	duration: 8.763s,	episode steps: 1678,	steps per second: 191,	episode reward: -19.000,	mean reward: -0.011 [-1.000, 1.000],	mean act
129944/1500000:	episode: 334,	duration: 10.006s,	episode steps: 1919,	steps per second: 192,	episode reward: -19.000,	mean reward: -0.010 [-1.000, 1.000],	mean act
132044/1500000:	episode: 335,	duration: 11.026s,	episode steps: 2100,	steps per second: 190,	episode reward: -18.000,	mean reward: -0.009 [-1.000, 1.000],	mean act
134210/1500000:	episode: 336,	duration: 11.387s,	episode steps: 2166,	steps per second: 190,	episode reward: -18.000,	mean reward: -0.008 [-1.000, 1.000],	mean act
136214/1500000:	episode: 337,	duration: 10.682s,	episode steps: 2004,	steps per second: 188,	episode reward: -18.000,	mean reward: -0.009 [-1.000, 1.000],	mean act
137742/1500000:	episode: 338,	duration: 7.969s,	episode steps: 1528,	steps per second: 192,	episode reward: -20.000,	mean reward: -0.013 [-1.000, 1.000],	mean act
139832/1500000:	episode: 339,	duration: 11.096s,	episode steps: 2090,	steps per second: 188,	episode reward: -16.000,	mean reward: -0.008 [-1.000, 1.000],	mean act
144201/1500000:	episode: 340,	duration: 11.713s,	episode steps: 2189,	steps per second: 187,	episode reward: -17.000,	mean reward: -0.008 [-1.000, 1.000],	mean act
144110/1500000:	episode: 341,	duration: 10.936s,	episode steps: 2089,	steps per second: 191,	episode reward: -16.000,	mean reward: -0.008 [-1.000, 1.000],	mean act
144964/1500000:	episode: 342,	duration: 14.842s,	episode steps: 2836,	steps per second: 191,	episode reward: -13.000,	mean reward: -0.005 [-1.000, 1.000],	mean act
148969/1500000:	episode: 343,	duration: 10.583s,	episode steps: 2023,	steps per second: 191,	episode reward: -18.000,	mean reward: -0.009 [-1.000, 1.000],	mean act
150986/1500000:	episode: 344,	duration: 10.801s,	episode steps: 2017,	steps per second: 187,	episode reward: -16.000,	mean reward: -0.008 [-1.000, 1.000],	mean act
154230/1500000:	episode: 345,	duration: 11.757s,	episode steps: 2244,	steps per second: 191,	episode reward: -18.000,	mean reward: -0.008 [-1.000, 1.000],	mean act
155397/1500000:	episode: 346,	duration: 11.333s,	episode steps: 2167,	steps per second: 191,	episode reward: -15.000,	mean reward: -0.007 [-1.000, 1.000],	mean act
159692/1500000:	episode: 347,	duration: 8.254s,	episode steps: 1585,	steps per second: 192,	episode reward: -19.000,	mean reward: -0.012 [-1.000, 1.000],	mean act
159394/1500000:	episode: 348,	duration: 12.675s,	episode steps: 2412,	steps per second: 190,	episode reward: -14.000,	mean reward: -0.006 [-1.000, 1.000],	mean act
161592/1500000:	episode: 349,	duration: 11.437s,	episode steps: 2198,	steps per second: 192,	episode reward: -15.000,	mean reward: -0.007 [-1.000, 1.000],	mean act
164256/1500000:	episode: 350,	duration: 13.947s,	episode steps: 2664,	steps per second: 191,	episode reward: -15.000,	mean reward: -0.006 [-1.000, 1.000],	mean act
166801/1500000:	episode: 351,	duration: 13.449s,	episode steps: 2545,	steps per second: 189,	episode reward: -14.000,	mean reward: -0.006 [-1.000, 1.000],	mean act
168886/1500000:	episode: 352,	duration: 10.877s,	episode steps: 2085,	steps per second: 192,	episode reward: -18.000,	mean reward: -0.009 [-1.000, 1.000],	mean act
170908/1500000:	episode: 353,	duration: 10.578s,	episode steps: 2022,	steps per second: 191,	episode reward: -20.000,	mean reward: -0.010 [-1.000, 1.000],	mean act
173473/1500000:	episode: 354,	duration: 13.366s,	episode steps: 2565,	steps per second: 192,	episode reward: -17.000,	mean reward: -0.007 [-1.000, 1.000],	mean act
176461/1500000:	episode: 355,	duration: 15.668s,	episode steps: 2988,	steps per second: 191,	episode reward: -16.000,	mean reward: -0.005 [-1.000, 1.000],	mean act
178949/1500000:	episode: 356,	duration: 12.947s,	episode steps: 2488,	steps per second: 192,	episode reward: -16.000,	mean reward: -0.006 [-1.000, 1.000],	mean act
181468/1500000:	episode: 357,	duration: 13.243s,	episode steps: 2519,	steps per second: 190,	episode reward: -15.000,	mean reward: -0.006 [-1.000, 1.000],	mean act
183568/1500000:	episode: 358,	duration: 11.061s,	episode steps: 2100,	steps per second: 190,	episode reward: -19.000,	mean reward: -0.009 [-1.000, 1.000],	mean act
186410/1500000:	episode: 359,	duration: 14.834s,	episode steps: 2842,	steps per second: 192,	episode reward: -14.000,	mean reward: -0.005 [-1	

```

672408/1500000: episode: 436, duration: 9.063s, episode steps: 1728, steps per second: 191, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean acti
674990/1500000: episode: 437, duration: 13.497s, episode steps: 2582, steps per second: 191, episode reward: -16.000, mean reward: -0.006 [-1.000, 1.000], mean acti
678180/1500000: episode: 438, duration: 17.025s, episode steps: 3190, steps per second: 187, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
680383/1500000: episode: 439, duration: 11.648s, episode steps: 2203, steps per second: 189, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
682959/1500000: episode: 440, duration: 13.668s, episode steps: 2576, steps per second: 188, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
685395/1500000: episode: 441, duration: 12.663s, episode steps: 2436, steps per second: 192, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
687015/1500000: episode: 442, duration: 8.466s, episode steps: 1620, steps per second: 191, episode reward: -20.000, mean reward: -0.012 [-1.000, 1.000], mean acti
689042/1500000: episode: 443, duration: 10.613s, episode steps: 2027, steps per second: 191, episode reward: -13.000, mean reward: -0.006 [-1.000, 1.000], mean acti
690772/1500000: episode: 444, duration: 9.039s, episode steps: 1730, steps per second: 191, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean acti
693458/1500000: episode: 445, duration: 14.070s, episode steps: 2686, steps per second: 191, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
695698/1500000: episode: 446, duration: 11.773s, episode steps: 2240, steps per second: 190, episode reward: -13.000, mean reward: -0.006 [-1.000, 1.000], mean acti
698096/1500000: episode: 447, duration: 12.565s, episode steps: 2398, steps per second: 191, episode reward: -11.000, mean reward: -0.005 [-1.000, 1.000], mean acti
700939/1500000: episode: 448, duration: 14.790s, episode steps: 2843, steps per second: 192, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
703196/1500000: episode: 449, duration: 11.770s, episode steps: 2257, steps per second: 192, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
705983/1500000: episode: 450, duration: 14.628s, episode steps: 2787, steps per second: 191, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
710973/1500000: episode: 451, duration: 10.362s, episode steps: 1990, steps per second: 192, episode reward: -14.000, mean reward: -0.007 [-1.000, 1.000], mean acti
710979/1500000: episode: 452, duration: 15.828s, episode steps: 3006, steps per second: 190, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
713594/1500000: episode: 453, duration: 13.658s, episode steps: 2615, steps per second: 191, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
715586/1500000: episode: 454, duration: 10.400s, episode steps: 1992, steps per second: 192, episode reward: -21.000, mean reward: -0.011 [-1.000, 0.000], mean acti
718358/1500000: episode: 455, duration: 14.485s, episode steps: 2772, steps per second: 191, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
721317/1500000: episode: 456, duration: 15.435s, episode steps: 2959, steps per second: 192, episode reward: -16.000, mean reward: -0.005 [-1.000, 1.000], mean acti
723305/1500000: episode: 457, duration: 10.455s, episode steps: 1988, steps per second: 190, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
724749/1500000: episode: 458, duration: 7.579s, episode steps: 1444, steps per second: 191, episode reward: -20.000, mean reward: -0.014 [-1.000, 1.000], mean acti
727313/1500000: episode: 459, duration: 13.501s, episode steps: 2564, steps per second: 190, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
729576/1500000: episode: 460, duration: 11.808s, episode steps: 2263, steps per second: 192, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
732247/1500000: episode: 461, duration: 13.970s, episode steps: 2671, steps per second: 191, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
735343/1500000: episode: 462, duration: 16.560s, episode steps: 3096, steps per second: 187, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
738095/1500000: episode: 463, duration: 14.457s, episode steps: 2752, steps per second: 190, episode reward: -17.000, mean reward: -0.006 [-1.000, 1.000], mean acti
740743/1500000: episode: 464, duration: 14.053s, episode steps: 2648, steps per second: 188, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
743965/1500000: episode: 465, duration: 16.888s, episode steps: 3222, steps per second: 191, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
746048/1500000: episode: 466, duration: 10.945s, episode steps: 2083, steps per second: 190, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
749412/1500000: episode: 467, duration: 17.589s, episode steps: 3364, steps per second: 191, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
751884/1500000: episode: 468, duration: 12.953s, episode steps: 2472, steps per second: 191, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
753792/1500000: episode: 469, duration: 9.987s, episode steps: 1908, steps per second: 191, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean acti
756161/1500000: episode: 470, duration: 12.495s, episode steps: 2369, steps per second: 190, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
758479/1500000: episode: 471, duration: 12.122s, episode steps: 2318, steps per second: 191, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
761294/1500000: episode: 472, duration: 14.705s, episode steps: 2815, steps per second: 191, episode reward: -18.000, mean reward: -0.006 [-1.000, 1.000], mean acti
764785/1500000: episode: 473, duration: 18.273s, episode steps: 3491, steps per second: 191, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
767757/1500000: episode: 474, duration: 15.493s, episode steps: 2972, steps per second: 192, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
770104/1500000: episode: 475, duration: 12.270s, episode steps: 2347, steps per second: 191, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
772356/1500000: episode: 476, duration: 11.917s, episode steps: 2252, steps per second: 189, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
774953/1500000: episode: 477, duration: 13.532s, episode steps: 2597, steps per second: 192, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
777146/1500000: episode: 478, duration: 11.396s, episode steps: 2193, steps per second: 192, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
779849/1500000: episode: 479, duration: 14.112s, episode steps: 2703, steps per second: 192, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
783048/1500000: episode: 480, duration: 16.670s, episode steps: 3199, steps per second: 192, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
785679/1500000: episode: 481, duration: 13.792s, episode steps: 2631, steps per second: 191, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
787971/1500000: episode: 482, duration: 12.035s, episode steps: 2292, steps per second: 190, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
790111/1500000: episode: 483, duration: 11.149s, episode steps: 2140, steps per second: 192, episode reward: -14.000, mean reward: -0.007 [-1.000, 1.000], mean acti
792795/1500000: episode: 484, duration: 13.986s, episode steps: 2684, steps per second: 192, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
795375/1500000: episode: 485, duration: 13.928s, episode steps: 2580, steps per second: 185, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
797860/1500000: episode: 486, duration: 12.999s, episode steps: 2485, steps per second: 191, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
800344/1500000: episode: 487, duration: 13.084s, episode steps: 2484, steps per second: 190, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
802728/1500000: episode: 488, duration: 12.590s, episode steps: 2384, steps per second: 189, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
804839/1500000: episode: 489, duration: 10.965s, episode steps: 2111, steps per second: 193, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
807532/1500000: episode: 490, duration: 14.069s, episode steps: 2693, steps per second: 191, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
809798/1500000: episode: 491, duration: 11.828s, episode steps: 2266, steps per second: 192, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
811915/1500000: episode: 492, duration: 11.051s, episode steps: 2117, steps per second: 192, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
814444/1500000: episode: 493, duration: 13.212s, episode steps: 2529, steps per second: 191, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
817025/1500000: episode: 494, duration: 13.487s, episode steps: 2581, steps per second: 191, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
819597/1500000: episode: 495, duration: 13.514s, episode steps: 2572, steps per second: 190, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
822773/1500000: episode: 496, duration: 16.529s, episode steps: 3176, steps per second: 192, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
825356/1500000: episode: 497, duration: 13.442s, episode steps: 2583, steps per second: 192, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
828108/1500000: episode: 498, duration: 14.383s, episode steps: 2752, steps per second: 191, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
831486/1500000: episode: 499, duration: 17.694s, episode steps: 3378, steps per second: 191, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
834978/1500000: episode: 500, duration: 18.361s, episode steps: 3492, steps per second: 190, episode reward: -12.000, mean reward: -0.003 [-1.000, 1.000], mean acti
837670/1500000: episode: 501, duration: 14.166s, episode steps: 2692, steps per second: 190, episode reward: -16.000, mean reward: -0.006 [-1.000, 1.000], mean acti
840526/1500000: episode: 502, duration: 14.893s, episode steps: 2856, steps per second: 192, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
844431/1500000: episode: 503, duration: 20.411s, episode steps: 3905, steps per second: 191, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
846368/1500000: episode: 504, duration: 10.121s, episode steps: 1937, steps per second: 191, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
848613/1500000: episode: 505, duration: 11.912s, episode steps: 2245, steps per second: 188, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
851422/1500000: episode: 506, duration: 14.621s, episode steps: 2809, steps per second: 192, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
854408/1500000: episode: 507, duration: 16.116s, episode steps: 2986, steps per second: 185, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
856791/1500000: episode: 508, duration: 12.584s, episode steps: 2383, steps per second: 189, episode reward: -20.000, mean reward: -0.008 [-1.000, 1.000], mean acti
859961/1500000: episode: 509, duration: 16.545s, episode steps: 3170, steps per second: 192, episode reward: -14.000, mean reward: -0.004 [-1.000, 1.000], mean acti
862878/1500000: episode: 510, duration: 15.333s, episode steps: 2917, steps per second: 190, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean acti
866154/1500000: episode: 511, duration: 17.201s, episode steps: 3276, steps per second: 190, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
868746/1500000: episode: 512, duration: 13.514s, episode steps: 2592, steps per second: 192, episode reward: -16.000, mean reward: -0.006 [-1.000, 1.000], mean acti
871054/1500000: episode: 513, duration: 12.095s, episode steps: 2308, steps per second: 191, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
874401/1500000: episode: 514, duration: 17.428s, episode steps: 3347, steps per second: 192, episode reward: -8.000, mean reward: -0.002 [-1.000, 1.000], mean acti
877193/1500000: episode: 515, duration: 14.662s, episode steps: 2792, steps per second: 190, episode reward: -16.000, mean reward: -0.006 [-1.000, 1.000], mean acti
880727/1500000: episode: 516, duration: 18.568s, episode steps: 3534, steps per second: 190, episode reward: -8.000, mean reward: -0.002 [-1.000, 1.000], mean acti
883971/1500000: episode: 517, duration: 16.922s, episode steps: 3244, steps per second: 192, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
885858/1500000: episode: 518, duration: 9.819s, episode steps: 1887, steps per second: 192, episode reward: -20.000, mean reward: -0.011 [-1.000, 1.000], mean acti
888289/1500000: episode: 519, duration: 12.705s, episode steps: 2431, steps per second: 191, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
891554/1500000: episode: 520, duration: 17.020s, episode steps: 3265, steps per second: 192, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean acti
895535/1500000: episode: 521, duration: 20.918s, episode steps: 3981, steps per second: 190, episode reward: -9.000, mean reward: -0.002 [-1.000, 1.000], mean acti
899015/1500000: episode: 522, duration: 18.096s, episode steps: 3480, steps per second: 192, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
902539/1500000: episode: 523, duration: 18.395s, episode steps: 3524, steps per second: 192, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
905394/1500000: episode: 524, duration: 14.872s, episode steps: 2855, steps per second: 192, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
908436/1500000: episode: 525, duration: 15.914s, episode steps: 3042, steps per second: 191, episode reward: -16.000, mean reward: -0.005 [-1.000, 1.000], mean acti
911845/1500000: episode: 526, duration: 18.094s, episode steps: 3409, steps per second: 188, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
915695/1500000: episode: 527, duration: 20.489s, episode steps: 3850, steps per second: 188, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean acti
918798/1500000: episode: 528, duration: 16.220s, episode steps: 3103, steps per second: 191, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
921474/1500000: episode: 529, duration: 13.957s, episode steps: 2676, steps per second: 192, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
924806/1500000: episode: 530, duration: 17.480s, episode steps: 3332, steps per second: 191, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
928269/1500000: episode: 531, duration: 17.999s, episode steps: 3463, steps per second: 192, episode reward: -14.000, mean reward: -0.004 [-1.000, 1.000], mean acti
931156/1500000: episode: 532, duration: 15.006s, episode steps: 2887, steps per second: 192, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean acti
934546/1500000: episode: 533, duration: 17.664s, episode steps: 3390, steps per second: 192, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
937031/1500000: episode: 534, duration: 12.913s, episode steps: 2485, steps per second: 192, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
940139/1500000: episode: 535, duration: 16.494s, episode steps: 3108, steps per second: 188, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
943183/1500000: episode: 536, duration: 15.860s, episode steps: 3044, steps per second: 192, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
946392/1500000: episode: 537, duration: 16.744s, episode steps: 3209, steps per second: 192, episode reward: -14.000, mean reward: -0.004 [-1.000, 1.000], mean acti
949417/1500000: episode: 538, duration: 15.737s, episode steps: 3025, steps per second: 192, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
952200/1500000: episode: 539, duration: 14.518s, episode steps: 2783, steps per second: 192, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean acti
954737/1500000: episode: 540, duration: 13.314s, episode steps: 2537, steps per second: 191, episode reward: -19.000, mean reward: -0.007 [-1.000, 1.000], mean acti
957265/1500000: episode: 541, duration: 13.253s, episode steps: 2528, steps per second: 191, episode reward: -18.000, mean reward: -0.007 [-1.000, 1.000], mean acti
960153/1500000: episode: 542, duration: 14.999s, episode steps: 2888, steps per second: 193, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
962505/1500000: episode: 543, duration: 12.274s, episode steps: 2352, steps per second: 192, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
965424/1500000: episode: 544, duration: 15.175s, episode steps: 2919, steps per second: 192, episode reward: -16.000, mean reward: -0.005 [-1.000, 1.000], mean acti

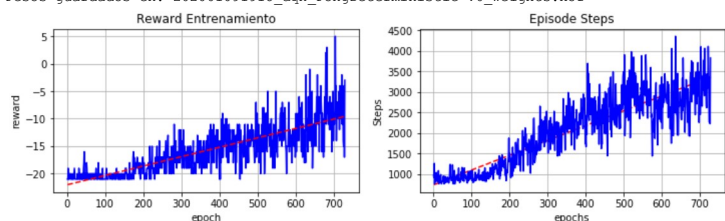
```

967961/1500000:	episode: 545,	duration: 13.241s,	episode steps: 2537,	steps per second: 192,	episode reward: -19.000,	mean reward: -0.007	[-1.000, 1.000],	mean acti
971068/1500000:	episode: 546,	duration: 16.601s,	episode steps: 3107,	steps per second: 187,	episode reward: -14.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
974120/1500000:	episode: 547,	duration: 16.264s,	episode steps: 3052,	steps per second: 188,	episode reward: -12.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
978002/1500000:	episode: 548,	duration: 20.245s,	episode steps: 3882,	steps per second: 192,	episode reward: -5.000,	mean reward: -0.001	[-1.000, 1.000],	mean acti
981023/1500000:	episode: 549,	duration: 15.748s,	episode steps: 3021,	steps per second: 192,	episode reward: -14.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
983502/1500000:	episode: 550,	duration: 12.879s,	episode steps: 2479,	steps per second: 192,	episode reward: -16.000,	mean reward: -0.006	[-1.000, 1.000],	mean acti
986088/1500000:	episode: 551,	duration: 13.564s,	episode steps: 2586,	steps per second: 191,	episode reward: -16.000,	mean reward: -0.006	[-1.000, 1.000],	mean acti
988509/1500000:	episode: 552,	duration: 12.573s,	episode steps: 2421,	steps per second: 193,	episode reward: -17.000,	mean reward: -0.007	[-1.000, 1.000],	mean acti
991740/1500000:	episode: 553,	duration: 16.802s,	episode steps: 3231,	steps per second: 192,	episode reward: -13.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
995015/1500000:	episode: 554,	duration: 17.024s,	episode steps: 3275,	steps per second: 192,	episode reward: -13.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
998032/1500000:	episode: 555,	duration: 15.702s,	episode steps: 3017,	steps per second: 192,	episode reward: -12.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1000961/1500000:	episode: 556,	duration: 15.372s,	episode steps: 2929,	steps per second: 191,	episode reward: -13.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1004210/1500000:	episode: 557,	duration: 16.933s,	episode steps: 3249,	steps per second: 192,	episode reward: -13.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1007115/1500000:	episode: 558,	duration: 15.127s,	episode steps: 2905,	steps per second: 192,	episode reward: -14.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1011019/1500000:	episode: 559,	duration: 20.307s,	episode steps: 3904,	steps per second: 192,	episode reward: -10.000,	mean reward: -0.003	[-1.000, 1.000],	mean acti
1014609/1500000:	episode: 560,	duration: 18.696s,	episode steps: 3590,	steps per second: 192,	episode reward: -9.000,	mean reward: -0.003	[-1.000, 1.000],	mean acti
1017296/1500000:	episode: 561,	duration: 14.142s,	episode steps: 2687,	steps per second: 190,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1020556/1500000:	episode: 562,	duration: 16.944s,	episode steps: 3260,	steps per second: 192,	episode reward: -8.000,	mean reward: -0.002	[-1.000, 1.000],	mean acti
1022947/1500000:	episode: 563,	duration: 12.454s,	episode steps: 2391,	steps per second: 192,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1025309/1500000:	episode: 564,	duration: 12.360s,	episode steps: 2362,	steps per second: 191,	episode reward: -15.000,	mean reward: -0.006	[-1.000, 1.000],	mean acti
1028319/1500000:	episode: 565,	duration: 15.640s,	episode steps: 3010,	steps per second: 192,	episode reward: -7.000,	mean reward: -0.002	[-1.000, 1.000],	mean acti
1031343/1500000:	episode: 566,	duration: 16.323s,	episode steps: 3024,	steps per second: 185,	episode reward: -11.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1034417/1500000:	episode: 567,	duration: 16.059s,	episode steps: 3074,	steps per second: 191,	episode reward: -9.000,	mean reward: -0.003	[-1.000, 1.000],	mean acti
1037081/1500000:	episode: 568,	duration: 13.878s,	episode steps: 2664,	steps per second: 192,	episode reward: -14.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1039933/1500000:	episode: 569,	duration: 14.866s,	episode steps: 2852,	steps per second: 192,	episode reward: -12.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1042862/1500000:	episode: 570,	duration: 15.280s,	episode steps: 2929,	steps per second: 192,	episode reward: -13.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1045960/1500000:	episode: 571,	duration: 16.190s,	episode steps: 3098,	steps per second: 191,	episode reward: -8.000,	mean reward: -0.003	[-1.000, 1.000],	mean acti
1049559/1500000:	episode: 572,	duration: 18.888s,	episode steps: 3599,	steps per second: 191,	episode reward: -8.000,	mean reward: -0.002	[-1.000, 1.000],	mean acti
1051949/1500000:	episode: 573,	duration: 12.426s,	episode steps: 2390,	steps per second: 192,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1054413/1500000:	episode: 574,	duration: 12.828s,	episode steps: 2464,	steps per second: 192,	episode reward: -15.000,	mean reward: -0.006	[-1.000, 1.000],	mean acti
1056348/1500000:	episode: 575,	duration: 10.116s,	episode steps: 1935,	steps per second: 191,	episode reward: -16.000,	mean reward: -0.008	[-1.000, 1.000],	mean acti
1057789/1500000:	episode: 576,	duration: 7.551s,	episode steps: 1441,	steps per second: 191,	episode reward: -20.000,	mean reward: -0.014	[-1.000, 1.000],	mean acti
1060040/1500000:	episode: 577,	duration: 11.797s,	episode steps: 2251,	steps per second: 191,	episode reward: -15.000,	mean reward: -0.007	[-1.000, 1.000],	mean acti
1062292/1500000:	episode: 578,	duration: 11.911s,	episode steps: 2252,	steps per second: 189,	episode reward: -15.000,	mean reward: -0.007	[-1.000, 1.000],	mean acti
1064643/1500000:	episode: 579,	duration: 12.315s,	episode steps: 2351,	steps per second: 191,	episode reward: -15.000,	mean reward: -0.006	[-1.000, 1.000],	mean acti
1066669/1500000:	episode: 580,	duration: 10.552s,	episode steps: 2026,	steps per second: 192,	episode reward: -15.000,	mean reward: -0.007	[-1.000, 1.000],	mean acti
1069579/1500000:	episode: 581,	duration: 15.190s,	episode steps: 2910,	steps per second: 192,	episode reward: -11.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1072262/1500000:	episode: 582,	duration: 14.015s,	episode steps: 2683,	steps per second: 191,	episode reward: -7.000,	mean reward: -0.003	[-1.000, 1.000],	mean acti
1074131/1500000:	episode: 583,	duration: 9.772s,	episode steps: 1869,	steps per second: 191,	episode reward: -16.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
1076404/1500000:	episode: 584,	duration: 11.963s,	episode steps: 2273,	steps per second: 190,	episode reward: -16.000,	mean reward: -0.007	[-1.000, 1.000],	mean acti
1079425/1500000:	episode: 585,	duration: 15.838s,	episode steps: 3021,	steps per second: 191,	episode reward: -12.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1082652/1500000:	episode: 586,	duration: 16.852s,	episode steps: 3227,	steps per second: 191,	episode reward: -13.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1085250/1500000:	episode: 587,	duration: 13.553s,	episode steps: 2598,	steps per second: 192,	episode reward: -15.000,	mean reward: -0.006	[-1.000, 1.000],	mean acti
1087921/1500000:	episode: 588,	duration: 13.937s,	episode steps: 2671,	steps per second: 192,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1090544/1500000:	episode: 589,	duration: 14.244s,	episode steps: 2623,	steps per second: 184,	episode reward: -14.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1092254/1500000:	episode: 590,	duration: 8.988s,	episode steps: 1710,	steps per second: 190,	episode reward: -18.000,	mean reward: -0.011	[-1.000, 1.000],	mean acti
1095058/1500000:	episode: 591,	duration: 14.721s,	episode steps: 2804,	steps per second: 190,	episode reward: -14.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1097723/1500000:	episode: 592,	duration: 13.927s,	episode steps: 2665,	steps per second: 191,	episode reward: -12.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1100512/1500000:	episode: 593,	duration: 14.502s,	episode steps: 2789,	steps per second: 192,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1103375/1500000:	episode: 594,	duration: 14.959s,	episode steps: 2863,	steps per second: 191,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1105712/1500000:	episode: 595,	duration: 12.186s,	episode steps: 2337,	steps per second: 192,	episode reward: -15.000,	mean reward: -0.006	[-1.000, 1.000],	mean acti
1108878/1500000:	episode: 596,	duration: 16.607s,	episode steps: 3166,	steps per second: 191,	episode reward: -13.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1112048/1500000:	episode: 597,	duration: 16.481s,	episode steps: 3170,	steps per second: 192,	episode reward: -12.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1114456/1500000:	episode: 598,	duration: 12.551s,	episode steps: 2408,	steps per second: 192,	episode reward: -15.000,	mean reward: -0.006	[-1.000, 1.000],	mean acti
1116725/1500000:	episode: 599,	duration: 11.848s,	episode steps: 2269,	steps per second: 192,	episode reward: -17.000,	mean reward: -0.007	[-1.000, 1.000],	mean acti
1119141/1500000:	episode: 600,	duration: 12.582s,	episode steps: 2416,	steps per second: 192,	episode reward: -16.000,	mean reward: -0.007	[-1.000, 1.000],	mean acti
1122508/1500000:	episode: 601,	duration: 17.630s,	episode steps: 3367,	steps per second: 191,	episode reward: -9.000,	mean reward: -0.003	[-1.000, 1.000],	mean acti
1124613/1500000:	episode: 602,	duration: 11.139s,	episode steps: 2105,	steps per second: 189,	episode reward: -16.000,	mean reward: -0.008	[-1.000, 1.000],	mean acti
1127785/1500000:	episode: 603,	duration: 16.603s,	episode steps: 3172,	steps per second: 191,	episode reward: -10.000,	mean reward: -0.003	[-1.000, 1.000],	mean acti
1130423/1500000:	episode: 604,	duration: 13.753s,	episode steps: 2638,	steps per second: 192,	episode reward: -11.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1133177/1500000:	episode: 605,	duration: 14.408s,	episode steps: 2754,	steps per second: 191,	episode reward: -14.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1136373/1500000:	episode: 606,	duration: 16.704s,	episode steps: 3196,	steps per second: 191,	episode reward: -9.000,	mean reward: -0.003	[-1.000, 1.000],	mean acti
1139115/1500000:	episode: 607,	duration: 14.452s,	episode steps: 2742,	steps per second: 190,	episode reward: -10.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1141933/1500000:	episode: 608,	duration: 14.763s,	episode steps: 2818,	steps per second: 191,	episode reward: -14.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1145378/1500000:	episode: 609,	duration: 17.908s,	episode steps: 3445,	steps per second: 192,	episode reward: -6.000,	mean reward: -0.002	[-1.000, 1.000],	mean acti
1148657/1500000:	episode: 610,	duration: 17.551s,	episode steps: 3279,	steps per second: 187,	episode reward: -9.000,	mean reward: -0.003	[-1.000, 1.000],	mean acti
1151563/1500000:	episode: 611,	duration: 15.201s,	episode steps: 2906,	steps per second: 191,	episode reward: -10.000,	mean reward: -0.003	[-1.000, 1.000],	mean acti
1154427/1500000:	episode: 612,	duration: 15.048s,	episode steps: 2864,	steps per second: 190,	episode reward: -11.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1157338/1500000:	episode: 613,	duration: 15.130s,	episode steps: 2911,	steps per second: 192,	episode reward: -8.000,	mean reward: -0.003	[-1.000, 1.000],	mean acti
1160146/1500000:	episode: 614,	duration: 14.636s,	episode steps: 2808,	steps per second: 192,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1162873/1500000:	episode: 615,	duration: 14.215s,	episode steps: 2727,	steps per second: 192,	episode reward: -16.000,	mean reward: -0.006	[-1.000, 1.000],	mean acti
1166126/1500000:	episode: 616,	duration: 16.978s,	episode steps: 3253,	steps per second: 192,	episode reward: -11.000,	mean reward: -0.003	[-1.000, 1.000],	mean acti
1169188/1500000:	episode: 617,	duration: 16.085s,	episode steps: 3062,	steps per second: 190,	episode reward: -7.000,	mean reward: -0.002	[-1.000, 1.000],	mean acti
1171568/1500000:	episode: 618,	duration: 12.472s,	episode steps: 2380,	steps per second: 191,	episode reward: -17.000,	mean reward: -0.007	[-1.000, 1.000],	mean acti
1174754/1500000:	episode: 619,	duration: 16.627s,	episode steps: 3186,	steps per second: 192,	episode reward: -12.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1177845/1500000:	episode: 620,	duration: 16.090s,	episode steps: 3091,	steps per second: 192,	episode reward: -11.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1180456/1500000:	episode: 621,	duration: 13.611s,	episode steps: 2611,	steps per second: 192,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1183088/1500000:	episode: 622,	duration: 13.785s,	episode steps: 2632,	steps per second: 191,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1185190/1500000:	episode: 623,	duration: 11.111s,	episode steps: 2102,	steps per second: 189,	episode reward: -18.000,	mean reward: -0.009	[-1.000, 1.000],	mean acti
1187761/1500000:	episode: 624,	duration: 13.367s,	episode steps: 2571,	steps per second: 192,	episode reward: -12.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1190384/1500000:	episode: 625,	duration: 13.654s,	episode steps: 2623,	steps per second: 192,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1192136/1500000:	episode: 626,	duration: 9.132s,	episode steps: 1752,	steps per second: 192,	episode reward: -18.000,	mean reward: -0.010	[-1.000, 1.000],	mean acti
1196192/1500000:	episode: 627,	duration: 21.145s,	episode steps: 4056,	steps per second: 192,	episode reward: -2.000,	mean reward: -0.000	[-1.000, 1.000],	mean acti
1199068/1500000:	episode: 628,	duration: 15.077s,	episode steps: 2876,	steps per second: 191,	episode reward: -12.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1201986/1500000:	episode: 629,	duration: 15.310s,	episode steps: 2918,	steps per second: 191,	episode reward: -11.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1204976/1500000:	episode: 630,	duration: 15.766s,	episode steps: 2990,	steps per second: 190,	episode reward: -10.000,	mean reward: -0.003	[-1.000, 1.000],	mean acti
1206913/1500000:	episode: 631,	duration: 10.307s,	episode steps: 1937,	steps per second: 188,	episode reward: -15.000,	mean reward: -0.008	[-1.000, 1.000],	mean acti
1209362/1500000:	episode: 632,	duration: 13.057s,	episode steps: 2449,	steps per second: 188,	episode reward: -16.000,	mean reward: -0.007	[-1.000, 1.000],	mean acti
1212649/1500000:	episode: 633,	duration: 17.133s,	episode steps: 3287,	steps per second: 192,	episode reward: -6.000,	mean reward: -0.002	[-1.000, 1.000],	mean acti
1215491/1500000:	episode: 634,	duration: 14.960s,	episode steps: 2842,	steps per second: 190,	episode reward: -11.000,	mean reward: -0.004	[-1.000, 1.000],	mean acti
1218003/1500000:	episode: 635,	duration: 13.095s,	episode steps: 2512,	steps per second: 192,	episode reward: -13.000,	mean reward: -0.005	[-1.000, 1.000],	mean acti
1220493/1500000:	episode: 636,	duration: 12.970s,	episode steps: 2490,	steps per second: 192,	episode reward: -14.000,	mean reward: -0.006	[-1.000, 1.000],	mean acti
1224851/1500000:	episode: 637,	duration: 22.718s,	episode steps: 4358,	steps per second: 192,	episode reward: -1.			

```

1269835/1500000: episode: 654, duration: 13.696s, episode steps: 2621, steps per second: 191, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1273171/1500000: episode: 655, duration: 17.471s, episode steps: 3336, steps per second: 191, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1275710/1500000: episode: 656, duration: 13.373s, episode steps: 2539, steps per second: 190, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1279541/1500000: episode: 657, duration: 20.084s, episode steps: 3831, steps per second: 191, episode reward: -2.000, mean reward: -0.001 [-1.000, 1.000], mean acti
1282732/1500000: episode: 658, duration: 16.704s, episode steps: 3191, steps per second: 191, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
1285389/1500000: episode: 659, duration: 13.883s, episode steps: 2657, steps per second: 191, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1288485/1500000: episode: 660, duration: 16.198s, episode steps: 3096, steps per second: 191, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean acti
1291512/1500000: episode: 661, duration: 15.945s, episode steps: 3027, steps per second: 190, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1293530/1500000: episode: 662, duration: 10.619s, episode steps: 2018, steps per second: 190, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
1296037/1500000: episode: 663, duration: 13.082s, episode steps: 2507, steps per second: 192, episode reward: -9.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1298757/1500000: episode: 664, duration: 14.237s, episode steps: 2720, steps per second: 191, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
1302059/1500000: episode: 665, duration: 17.269s, episode steps: 3302, steps per second: 191, episode reward: -8.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1305325/1500000: episode: 666, duration: 17.083s, episode steps: 3266, steps per second: 191, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean acti
1308185/1500000: episode: 667, duration: 15.103s, episode steps: 2860, steps per second: 189, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1311194/1500000: episode: 668, duration: 15.730s, episode steps: 3009, steps per second: 191, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
1313454/1500000: episode: 669, duration: 11.818s, episode steps: 2260, steps per second: 191, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
1316528/1500000: episode: 670, duration: 16.124s, episode steps: 3074, steps per second: 191, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
1319356/1500000: episode: 671, duration: 14.767s, episode steps: 2828, steps per second: 192, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
1321599/1500000: episode: 672, duration: 12.049s, episode steps: 2243, steps per second: 186, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
1323529/1500000: episode: 673, duration: 10.157s, episode steps: 1930, steps per second: 190, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
1326272/1500000: episode: 674, duration: 14.853s, episode steps: 2743, steps per second: 185, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1328201/1500000: episode: 675, duration: 10.117s, episode steps: 1929, steps per second: 191, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
1331563/1500000: episode: 676, duration: 17.527s, episode steps: 3362, steps per second: 192, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1335032/1500000: episode: 677, duration: 18.192s, episode steps: 3469, steps per second: 191, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1338305/1500000: episode: 678, duration: 17.324s, episode steps: 3273, steps per second: 189, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1342062/1500000: episode: 679, duration: 19.713s, episode steps: 3757, steps per second: 191, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1344822/1500000: episode: 680, duration: 14.479s, episode steps: 2760, steps per second: 191, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1347949/1500000: episode: 681, duration: 16.443s, episode steps: 3127, steps per second: 190, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
1351668/1500000: episode: 682, duration: 19.722s, episode steps: 3719, steps per second: 189, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1354599/1500000: episode: 683, duration: 17.494s, episode steps: 3291, steps per second: 188, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1357907/1500000: episode: 684, duration: 15.661s, episode steps: 2948, steps per second: 188, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
1361238/1500000: episode: 685, duration: 17.673s, episode steps: 3331, steps per second: 188, episode reward: -8.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1364487/1500000: episode: 686, duration: 17.320s, episode steps: 3249, steps per second: 188, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
1367826/1500000: episode: 687, duration: 17.933s, episode steps: 3339, steps per second: 186, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1370716/1500000: episode: 688, duration: 15.470s, episode steps: 2890, steps per second: 187, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
1373162/1500000: episode: 689, duration: 13.073s, episode steps: 2446, steps per second: 187, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
1376356/1500000: episode: 690, duration: 17.094s, episode steps: 3194, steps per second: 187, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1379849/1500000: episode: 691, duration: 18.588s, episode steps: 3487, steps per second: 185, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
1382897/1500000: episode: 692, duration: 16.781s, episode steps: 3054, steps per second: 182, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1385607/1500000: episode: 693, duration: 14.941s, episode steps: 2710, steps per second: 181, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1389039/1500000: episode: 694, duration: 18.556s, episode steps: 3432, steps per second: 185, episode reward: -8.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1392133/1500000: episode: 695, duration: 16.779s, episode steps: 3094, steps per second: 184, episode reward: -16.000, mean reward: -0.005 [-1.000, 1.000], mean acti
1395849/1500000: episode: 696, duration: 20.235s, episode steps: 3716, steps per second: 184, episode reward: -9.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1398757/1500000: episode: 697, duration: 16.019s, episode steps: 2908, steps per second: 182, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1402256/1500000: episode: 698, duration: 19.059s, episode steps: 3499, steps per second: 184, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean acti
1404513/1500000: episode: 699, duration: 12.399s, episode steps: 2257, steps per second: 182, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
1407397/1500000: episode: 700, duration: 15.725s, episode steps: 2884, steps per second: 183, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
1410090/1500000: episode: 701, duration: 14.763s, episode steps: 2693, steps per second: 182, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1412542/1500000: episode: 702, duration: 13.544s, episode steps: 2452, steps per second: 181, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1415217/1500000: episode: 703, duration: 14.589s, episode steps: 2675, steps per second: 183, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean acti
1419321/1500000: episode: 704, duration: 22.499s, episode steps: 4104, steps per second: 182, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1422627/1500000: episode: 705, duration: 18.052s, episode steps: 3306, steps per second: 183, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1425966/1500000: episode: 706, duration: 18.286s, episode steps: 3339, steps per second: 183, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean acti
1428621/1500000: episode: 707, duration: 14.498s, episode steps: 2655, steps per second: 183, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
1431796/1500000: episode: 708, duration: 17.342s, episode steps: 3175, steps per second: 183, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1435185/1500000: episode: 709, duration: 18.656s, episode steps: 3389, steps per second: 182, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1438226/1500000: episode: 710, duration: 16.602s, episode steps: 3041, steps per second: 183, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean acti
1442002/1500000: episode: 711, duration: 21.159s, episode steps: 3776, steps per second: 178, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1445218/1500000: episode: 712, duration: 17.669s, episode steps: 3216, steps per second: 182, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
1448128/1500000: episode: 713, duration: 15.959s, episode steps: 2910, steps per second: 182, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean acti
1452191/1500000: episode: 714, duration: 22.272s, episode steps: 4063, steps per second: 182, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean acti
1455196/1500000: episode: 715, duration: 16.659s, episode steps: 3005, steps per second: 180, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1458642/1500000: episode: 716, duration: 19.054s, episode steps: 3446, steps per second: 181, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean acti
1461707/1500000: episode: 717, duration: 16.906s, episode steps: 3065, steps per second: 181, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1464368/1500000: episode: 718, duration: 14.727s, episode steps: 2661, steps per second: 181, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1466983/1500000: episode: 719, duration: 14.497s, episode steps: 2615, steps per second: 180, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
1470268/1500000: episode: 720, duration: 18.287s, episode steps: 3285, steps per second: 180, episode reward: -8.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1473699/1500000: episode: 721, duration: 18.973s, episode steps: 3431, steps per second: 181, episode reward: -2.000, mean reward: -0.001 [-1.000, 1.000], mean acti
1476798/1500000: episode: 722, duration: 17.180s, episode steps: 3099, steps per second: 180, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1480908/1500000: episode: 723, duration: 22.699s, episode steps: 4110, steps per second: 181, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1484507/1500000: episode: 724, duration: 20.021s, episode steps: 3599, steps per second: 180, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
1487201/1500000: episode: 725, duration: 14.957s, episode steps: 2694, steps per second: 180, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
1489466/1500000: episode: 726, duration: 12.746s, episode steps: 2265, steps per second: 178, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
1492829/1500000: episode: 727, duration: 18.645s, episode steps: 3363, steps per second: 180, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean acti
1495055/1500000: episode: 728, duration: 12.542s, episode steps: 2226, steps per second: 177, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
1498883/1500000: episode: 729, duration: 21.675s, episode steps: 3828, steps per second: 177, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean acti
done, took 7925.187 seconds
Pesos guardados en: 20201091918_dqn_PongDeterministic-v0_weights.h5f

```



Testing for 10 episodes ...

Episode 1: reward: 9.000, steps: 3027

Episode 2: reward: 11.000, steps: 2772

Episode 3: reward: -8.000, steps: 2788

Episode 4: reward: -4.000, steps: 3246

Episode 5: reward: 2.000, steps: 4033

Episode 6: reward: -9.000, steps: 2847

Episode 7: reward: -14.000, steps: 2920

Episode 8: reward: -8.000, steps: 2876

Episode 9: reward: -9.000, steps: 2655

Episode 10: reward: -6.000, steps: 3206

mean reward: -3.6

mean_steps: 3037

(-3.6, 3037)