```
972/500000: episode: 1, duration: 3.933s, episode steps: 972, steps per second: 247, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 2.
 1886/500000: episode: 2,
                                duration: 2.819s, episode steps: 914, steps per second: 324, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action: 2.
 2650/500000: episode: 3,
                                duration: 2.384s, episode steps: 764, steps per second: 320, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.384s, episode steps: 764, steps per second: 320, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.384s, episode steps: 764, steps per second: 320, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.384s, episode steps: 764, steps per second: 320, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.384s, episode steps: 764, steps per second: 320, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.384s, episode steps: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.384s, episode steps: -21.000, episode steps: -21.0
                                                       episode steps: 842, steps per second: 315, episode reward: -21.000, mean reward: -0.025
                                                                                                                                                                                   0.000], mean action:
 4376/500000: episode: 5,
                                duration: 2.824s, episode steps: 884, steps per second: 313, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2.
                                duration: 3.149s, episode steps: 975, steps per second: 310, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 2.
 6282/500000: episode: 7,
                                duration: 2.823s, episode steps: 931, steps per second: 330, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action: 2.
                                                       episode steps: 997, steps per second: 329, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action: 2
  7279/500000
                                duration: 3.032s,
 8317/500000: episode: 9, duration: 3.129s, episode steps: 1038, steps per second: 332, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action: 2
                 episode: 10, duration: 2.822s, episode steps: 934, steps per second: 331, episode reward: -21.000, mean reward: -0.022
10135/500000: episode: 11, duration: 5.712s, episode steps: 884, steps per second: 155, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2
11050/500000:
                 episode: 12, duration: 16.478s, episode steps: 915, steps per second: 56, episode reward: -20.000, mean reward: -0.022 [-1.000,
                                                                                                                                                                                   1.000], mean action: 2
12027/500000: episode: 13, duration: 17.601s, episode steps: 977, steps per second: 56, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action: 2
                 episode: 14, duration: 15.671s, episode steps: 880, steps per second: 56, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2
13699/500000: episode: 15, duration: 14.130s, episode steps: 792, steps per second: 56, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2
                 episode: 16, duration: 18.072s, episode steps: 1019, steps per second: 56, episode reward: -21.000, mean reward: -0.021
14718/500000:
15789/500000: episode: 17, duration: 18.823s, episode steps: 1071, steps per second: 57, episode reward: -19.000, mean reward: -0.018 [-1.000, 1.000], mean action:
                 episode: 18, duration: 14.950s, episode steps: 842, steps per second: 56, episode reward: -21.000, mean reward: -0.025 [-1.000, 0
                                                                                                                                                                                      .000], mean action: 2
17474/500000: episode: 19, duration: 14.965s, episode steps: 843, steps per second: 56, episode reward: -21.000, mean reward: -0.025 (-1.000, 0.000), mean action: 2
                 episode: 20, duration: 17.505s, episode steps: 1003, steps per second: 57, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action:
18477/500000:
19349/500000: episode: 21, duration: 15.563s, episode steps: 872, steps per second: 56, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 20160/500000: episode: 22, duration: 14.346s, episode steps: 811, steps per second: 57, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 20160/500000: episode: 22, duration: 14.346s, episode steps: 811, steps per second: 57, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 20160/500000: episode: 22, duration: 14.346s, episode steps: 811, steps per second: 57, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 20160/500000: episode: 22, duration: 14.346s, episode steps: 811, steps per second: 57, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 20160/500000: episode: 22, duration: 14.346s, episode steps: 811, steps per second: 57, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 20160/500000: episode: 22, duration: 14.346s, episode steps: 811, steps per second: 57, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 20160/500000: episode: 22, duration: 14.346s, episode steps: 811, steps per second: 57, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 20160/500000: episode: 20160/50000000: episode: 20160/5000000: episode: 20160/500000: episode: 201
21013/500000: episode: 23, duration: 14.849s, episode steps: 853, steps per second: 57, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2
                 episode: 24, duration: 15.776s, episode steps: 898, steps per second: 57, episode reward: -20.000, mean reward: -0.022
22735/500000: episode: 25, duration: 14.532s, episode steps: 824, steps per second: 57, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2
                 episode: 26, duration: 16.076s, episode steps: 903, steps per second: 56, episode reward: -20.000, mean reward: -0.022
24542/500000: episode: 27, duration: 15.969s, episode steps: 904, steps per second: 57, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action: 2
                 episode: 28, duration: 14.521s, episode steps: 824, steps per second: 57, episode reward: -21.000, mean reward:
26130/500000: episode: 29, duration: 13.453s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                                                         [-1.000, 0.000], mean action: 2
                 episode: 30, duration: 14.992s, episode steps: 826, steps per second: 55, episode reward: -21.000, mean reward: -0.025
27826/500000: episode: 31, duration: 15.319s, episode steps: 870, steps per second: 57, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2
28770/500000:
                 episode: 32, duration: 16.625s, episode steps: 944, steps per second: 57, episode reward: -21.000, mean reward: -0.022
29736/500000: episode: 33, duration: 16.934s, episode steps: 966, steps per second: 57, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action: 2
30700/500000:
                 episode: 34, duration: 17.196s, episode steps: 964, steps per second: 56, episode reward: -20.000, mean reward: -0.021
31643/500000: episode: 35, duration: 16.462s, episode steps: 943, steps per second: 57, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 2
                  episode: 36, duration: 15.042s, episode steps: 854,
                                                                                   steps per second: 57, episode reward: -21.000, mean reward: -0.025 [-1.000,
32497/500000.
33454/500000: episode: 37, duration: 16.807s, episode steps: 957, steps per second: 57, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action: 2
                 episode: 38, duration: 17.789s, episode steps: 1011, steps per second:
                                                                                                           57, episode reward: -19.000, mean reward: -0.019 [-1.000,
35291/500000: episode: 39, duration: 14.405s, episode steps: 826, steps per second: 57, episode reward: -21.000, mean reward: -0.025 (-1.000, 0.000), mean action: 2
                 episode: 40, duration: 15.841s, episode steps: 882, steps per second: 56, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2
36173/500000:
37135/500000: episode: 41, duration: 16.842s, episode steps: 962, steps per second: 57, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action: 2
38221/500000:
                 episode: 42, duration: 19.141s, episode steps: 1086, steps per second: 57,
                                                                                                                 episode reward: -19.000, mean reward: -0.017 [-1.000,
39152/500000: episode: 43, duration: 16.362s, episode steps: 931, steps per second: 57, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action: 2
                                                          episode steps: 974, steps per second: 57, episode reward: -21.000, mean reward: -0.022 [-1.000,
40126/500000
40938/500000: episode: 45, duration: 14.237s, episode steps: 812, steps per second: 57, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 2
41944/500000:
                 episode: 46, duration: 17.736s, episode steps: 1006, steps per second: 57,
                                                                                                                episode reward: -20.000, mean reward: -0.020 [-1.000,
                                                                                                                                                                                     1.000], mean action:
42924/500000; episode: 47, duration: 17.385s, episode steps: 980, steps per second: 56, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action: 2
                 episode: 48, duration: 16.636s, episode steps: 938, steps per second: 56, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action: 2
43862/500000:
44808/500000: episode: 49, duration: 17.132s, episode steps: 946, steps per second: 55, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action:
45632/500000:
                 episode: 50, duration: 14.466s, episode steps: 824, steps per second: 57, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2
46475/500000: episode: 51, duration: 14.919s, episode steps: 843, steps per second: 57, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2
                 episode: 52, duration: 14.307s, episode steps: 820,
                                                                                   steps per second: 57, episode reward: -21.000, mean reward: -0.026
47295/500000:
48272/500000: episode: 53, duration: 17.338s, episode steps: 977, steps per second: 56, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action: 2
49098/500000:
                 episode: 54, duration: 14.580s, episode steps: 826, steps per second: 57, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2
50239/500000: episode: 55, duration: 19.833s, episode steps: 1141, steps per second: 58, episode reward: -18.000, mean reward: -0.016 [-1.000, 1.000], mean action:
51172/500000:
                  episode: 56, duration: 16.275s, episode steps: 933, steps per second: 57, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action: 2
52115/500000: episode: 57, duration: 16.461s, episode steps: 943, steps per second: 57, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action: 2
                 episode: 58, duration: 16.783s, episode steps:
                                                                             965, steps per second: 57, episode reward: -21.000, mean reward: -0.022 [-1.000,
53872/500000: episode: 59, duration: 14.007s, episode steps: 792, steps per second: 57, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2
54953/500000:
                 episode: 60, duration: 18.826s, episode steps: 1081, steps per second: 57, episode reward: -19.000, mean reward: -0.018 [-1.000,
                                                                                                                                                                                     1.000], mean action:
55940/500000: episode: 61, duration: 17.179s, episode steps: 987, steps per second: 57, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action: 2
                 episode: 62, duration: 19.328s, episode steps: 1109, steps per second: 57,
                                                                                                                episode reward: -20.000, mean reward: -0.018 [-1.000, 1.000], mean action:
57920/500000: episode: 63, duration: 15.216s, episode steps: 871, steps per second: 57, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2
                 episode: 64, duration: 19.060s, episode steps: 1099, steps per second: 58, episode reward: -18.000, mean reward: -0.016 [-1.000,
59019/500000.
                                                                                                                                                                                     1.0001, mean action:
59783/500000: episode: 65, duration: 13.227s, episode steps: 764, steps per second: 58, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2
                 episode: 66, duration: 20.733s, episode steps:
                                                                             1200, steps per second: 58, episode reward: -17.000, mean reward: -0.014 [-1.000,
61835/500000: episode: 67, duration: 14.964s, episode steps: 852, steps per second: 57, episode reward: -21.000, mean reward: -0.025 (-1.000, 0.000), mean action: 2
                 episode: 68, duration: 15.904s, episode steps: 902, steps per second: 57, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 2
62737/500000:
63636/500000: episode: 69, duration: 15.403s, episode steps: 899, steps per second: 58, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action: 2
                  episode: 70, duration: 17.642s, episode steps: 1016, steps per second: 58,
                                                                                                                 episode reward: -19.000, mean reward: -0.019 [-1.000,
64652/500000:
65494/500000: episode: 71, duration: 14.554s, episode steps: 842, steps per second: 58, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2
                 episode: 72, duration: 19.231s, episode steps: 1115, steps per second: 58, episode reward: -18.000, mean reward: -0.016 [-1.000,
66609/500000:
                                                                                                                                                                                     1.000], mean action:
67433/500000: episode: 73, duration: 14.353s, episode steps: 824, steps per second: 57, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 268197/500000: episode: 74, duration: 13.313s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2
69194/500000: episode: 75, duration: 17.537s, episode steps: 997, steps per second: 57, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action: 2
                 episode: 76, duration: 17.908s, episode steps: 1029, steps per second: 57, episode reward: -21.000, mean reward: -0.020 [-1.000,
70223/500000:
70987/500000: episode: 77, duration: 13.526s, episode steps: 764, steps per second: 56, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2
                 episode: 78, duration: 16.765s, episode steps: 958, steps per second: 57, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action: 2
71945/500000.
72832/500000: episode: 79, duration: 15.354s, episode steps: 887, steps per second: 58, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 2
                 episode: 80, duration: 14.804s, episode steps:
                                                                             843, steps per second: 57, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2
73675/500000:
74735/500000: episode: 81, duration: 18.396s, episode steps: 1060, steps per second: 58, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action:
75653/500000:
                 episode: 82, duration: 15.966s, episode steps: 918, steps per second: 57, episode reward: -20.000, mean reward: -0.022 [-1.000, 1
                                                                                                                                                                                      .000], mean action: 2
76445/500000: episode: 83, duration: 13.796s, episode steps: 792, steps per second: 57, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2
                 episode: 84, duration: 14.810s, episode steps: 842,
                                                                                   steps per second: 57, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2
77287/500000
78200/500000: episode: 85, duration: 16.115s, episode steps: 913, steps per second: 57, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action: 2
                 episode: 86, duration: 14.501s, episode steps:
                                                                                   steps per second: 57, episode reward: -21.000, mean reward: -0.025
79025/500000.
                                                                                                                                                                         [-1.000, 0.000], mean action:
79897/500000: episode: 87, duration: 15.882s, episode steps: 872, steps per second: 55, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2
80661/500000:
                 episode: 88, duration: 13.609s, episode steps: 764, steps per second: 56, episode reward: -21.000, mean reward: -0.027
81545/500000:
                 episode: 89, duration: 15.350s, episode steps: 884, steps per second: 58, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
                 episode: 90, duration: 16.962s, episode steps: 972, steps per second: 57, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 2
82517/500000.
83430/500000: episode: 91, duration: 15.891s, episode steps: 913, steps per second: 57, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action: 2
                 episode: 92, duration: 16.731s, episode steps: 963,
                                                                                   steps per second: 58, episode reward: -21.000, mean reward: -0.022
84393/500000.
                                                                                                                                                                          -1.000,
85157/500000: episode: 93, duration: 13.328s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                                                         [-1.000, 0.000], mean action: 2
                 episode: 94, duration: 16.323s, episode steps: 945, steps per second: 58, episode reward: -21.000, mean reward: -0.022
86102/500000:
                                                                                                                                                                         [-1.000, 0.000], mean action:
86984/500000: episode: 95, duration: 15.378s, episode steps: 882, steps per second: 57, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2
87836/500000:
                 episode: 96, duration: 14.904s, episode steps: 852, steps per second: 57, episode reward: -21.000, mean reward: -0.025
88688/500000: episode: 97, duration: 15.072s, episode steps: 852, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                                                        [-1.000, 0.000], mean action: 2
89574/500000: episode: 98, duration: 15.405s, episode steps: 886, steps per second: 58, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2
90338/500000: episode: 99, duration: 13.479s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 291102/500000: episode: 100, duration: 13.310s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
91956/500000: episode: 101, duration: 14.991s, episode steps: 854, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                                                          [-1.000, 0.000], mean action:
92748/500000: episode: 102, duration: 14.046s, episode steps: 792, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                                                          [-1.000, 0.000], mean action:
93572/500000: episode: 103. duration: 14.565s, episode steps: 824. steps per second: 57. episode reward: -21.000. mean reward: -0.025
                                                                                                                                                                          [-1.000, 0.000], mean action:
                                                                                                                episode reward: -20.000,
                                                                                                                                                                          [-1.000,
94414/500000.
                 episode: 104, duration: 14.830s, episode steps: 842,
                                                                                            per second: 57,
                                                                                                                                                mean reward: -0.024
                                                                                                                                                                                     1.000], mean action:
95224/500000: episode: 105, duration: 14.313s, episode steps: 810, steps per second: 57, episode reward: -21.000, mean reward: -0.026
                                                                                                                                                                          [-1.000, 0.000], mean action:
96048/500000: episode: 106, duration: 14.619s, episode steps: 824, steps per second: 56, episode reward: -21.000, mean reward: -0.025 96812/500000: episode: 107, duration: 13.605s, episode steps: 764, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                                                          [-1.000, 0.000], mean action:
                                                                                                                                                                          [-1.000, 0.000], mean action:
97636/500000: episode: 108, duration: 15.035s, episode steps: 824, steps per second: 55, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
```

```
98488/500000: episode: 109, duration: 14.925s, episode steps: 852, steps per second: 57, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
99280/500000: episode: 110, duration: 13.692s, episode steps: 792, steps per second: 58, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
                                                                764, steps per second: 57, episode reward: -21.000, mean reward: -0.027
               episode: 111, duration: 13.519s, episode steps:
100836/500000: episode: 112, duration: 13.915s, episode steps: 792, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000, 0.000], mean action:
101600/500000: episode: 113, duration: 13.392s, episode steps:
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
                                                                 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027
102424/500000: episode: 114, duration: 14.543s, episode steps: 824, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
               episode: 115, duration:
                                       17.015s, episode steps: 973, steps per second:
                                                                                        57, episode reward: -21.000, mean reward: -0.022
                                                                                                                                                    0.000], mean action:
104161/500000: episode: 116, duration: 13.371s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
               episode: 117, duration: 14.669s, episode steps:
                                                                                        57, episode reward: -20.000, mean reward: -0.024
                                                                842, steps per second:
                                                                                                                                                     .000], mean action:
105795/500000: episode: 118, duration: 14.162s, episode steps: 792, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000, 0.000], mean action:
               episode: 119, duration: 13.370s, episode steps:
                                                                 764, steps per second: 57, episode reward: -21.000, mean reward:
                                                                                                                                             -1.000,
                                                                                                                                                    0.000], mean action:
107323/500000: episode: 120, duration: 13.428s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
               episode: 121, duration: 13.276s, episode steps: 764, steps per second: 58,
                                                                                            episode reward: -21.000, mean reward: -0.027
                                                                                                                                                    0.000], mean action:
108851/500000: episode: 122, duration: 13.516s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000, 0.000], mean action:
               episode: 123, duration:
                                        14.492s, episode steps:
                                                                826, steps per second:
                                                                                            episode reward: -21.000, mean reward: -0.025
                                                                                                                                                    0.000], mean action:
110441/500000: episode: 124, duration: 13.346s, episode steps: 764, steps per second: 57,
                                                                                            episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
               episode: 125, duration: 13.326s, episode steps:
                                                                764, steps per second: 57,
                                                                                            episode reward: -21.000, mean reward: -0.027
                                                                                                                                             -1.000,
                                                                                                                                                    0.000], mean action:
112029/500000: episode: 126, duration: 14.591s, episode steps: 824, steps per second: 56, episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000,
                                                                                                                                                    0.0001, mean action:
                                                                                                                                           [-1.000,
                                                                                                                                                    1.000], mean action:
               episode: 127, duration:
                                       15.680s, episode steps: 898, steps per second: 57, episode reward: -20.000, mean reward: -0.022
113797/500000: episode: 128, duration: 15.299s, episode steps: 870, steps per second: 57, episode reward: -20.000, mean reward: -0.023
                                                                                                                                           [-1.000,
                                                                                                                                                    1.000], mean action:
114809/500000: episode: 129, duration:
                                        17.905s, episode steps: 1012, steps per second: 57, episode reward: -19.000, mean reward: -
                                                                                                                                     -0.019 [-1.000,
                                                                                                                                                     1.000], mean action:
115573/500000: episode: 130, duration: 13.471s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 116614/500000: episode: 131, duration: 17.954s, episode steps: 1041, steps per second: 58, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action:
117426/500000: episode: 132, duration: 14.004s, episode steps: 812, steps per second: 58, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action:
118190/500000: episode: 133, duration: 13.489s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
118954/500000: episode: 134, duration: 13.414s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000,
                                                                                                                                                    0.0001, mean action:
                                       13.381s, episode steps:
               episode: 135, duration:
                                                                764, steps per second:
                                                                                            episode reward: -21.000, mean reward: -0.027
120482/500000: episode: 136. duration: 13.460s. episode steps: 764. steps per second: 57. episode reward: -21.000. mean reward: -0.027
                                                                                                                                           [-1.000, 0.000], mean action:
               episode: 137, duration: 13.339s, episode steps: 764, steps per second:
                                                                                        57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                                    0.000], mean action:
122088/500000; episode: 138, duration: 14.893s, episode steps: 842, steps per second: 57, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000.
                                                                                                                                                    1.0001, mean action:
               episode: 139, duration:
                                       16.545s, episode steps: 931, steps per second: 56, episode reward: -20.000, mean reward: -0.021
                                                                                                                                           [-1.000,
                                                                                                                                                    1.000], mean action:
123783/500000: episode: 140, duration: 13.646s, episode steps: 764, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
                                        13.605s, episode steps:
                                                                764, steps per second: 56,
125311/500000: episode: 142, duration: 13.524s, episode steps: 764, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000, 0.000], mean action:
               episode: 143, duration:
                                        13.302s, episode steps:
                                                                764, steps per second:
                                                                                            episode reward: -21.000, mean reward: -0.027
                                                                                                                                                    0.000], mean action:
126839/500000; episode: 144, duration: 13.527s, episode steps: 764, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
127650/500000:
               episode: 145, duration: 14.287s, episode steps: 811, steps per second: 57, episode reward: -21.000, mean reward: -0.026
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
128414/500000: episode: 146, duration: 13.462s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
                                                                764,
               episode: 147, duration:
                                       13.508s, episode steps:
129178/500000:
                                                                     steps per second: 57,
                                                                                            episode reward: -21.000,
                                                                                                                      mean reward: -0.027
                                                                                                                                             -1.000,
                                                                                                                                                    0.000], mean action:
129942/500000: episode: 148, duration: 13.446s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
               episode: 149, duration: 15.336s, episode steps: 871, steps per second: 57,
                                                                                            episode reward: -20.000, mean reward: -0.023
                                                                                                                                                    1.000], mean action:
131758/500000: episode: 150. duration: 16.727s. episode steps: 945. steps per second: 56. episode reward: -19.000. mean reward: -0.020
                                                                                                                                           [-1.000.
                                                                                                                                                    1.0001, mean action:
               episode: 151, duration:
132578/500000:
                                       14.694s, episode steps:
                                                                820, steps per second: 56,
                                                                                            episode reward: -21.000, mean reward: -0.026
                                                                                                                                            -1.000,
                                                                                                                                                    0.000], mean action:
133403/500000; episode: 152, duration: 14.461s, episode steps: 825, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
               episode: 153, duration:
                                       15.406s, episode steps: 870, steps per second: 56,
                                                                                            episode reward: -20.000,
                                                                                                                      mean reward: -0.023
                                                                                                                                                    1.000], mean action:
135155/500000: episode: 154, duration: 15.690s, episode steps: 882, steps per second: 56, episode reward: -21.000, mean reward: -0.024
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
               episode: 155, duration:
                                        13.854s, episode steps:
                                                                     steps per second:
                                                                                            episode reward: -21.000,
                                                                                                                      mean reward: -0.027
                                                                                                                                             -1.000,
                                                                                                                                                    0.000], mean action:
136739/500000: episode: 156, duration: 14.099s, episode steps: 792, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
               episode: 157, duration: 14.192s, episode steps: 811, steps per second:
                                                                                        57, episode reward: -21.000, mean reward: -0.026
                                                                                                                                                    0.000], mean action:
138342/500000; episode: 158, duration: 13.937s, episode steps: 792, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
139106/500000:
               episode: 159, duration:
                                       13.454s, episode steps:
                                                                764, steps per second: 57,
                                                                                            episode reward: -21.000, mean reward: -0.027
                                                                                                                                            -1.000,
                                                                                                                                                    0.000], mean action:
139870/500000: episode: 160, duration: 13.485s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
               episode: 161, duration: 13.692s, episode steps:
                                                                764, steps per second: 56, episode reward: -21.000, mean reward:
                                                                                                                                             1.000,
                                                                                                                                                    0.000], mean action:
141472/500000: episode: 162, duration: 14.856s, episode steps: 838, steps per second: 56, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000,
                                                                                                                                                    1.000], mean action:
                                                                     steps per second: 56, episode reward: -19.000,
               episode: 163, duration:
                                       16.809s, episode steps:
                                                                                                                      mean reward: -0.020
                                                                                                                                                    1.000], mean action:
143184/500000: episode: 164. duration: 13.582s. episode steps: 764. steps per second: 56. episode reward: -21.000. mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
               episode: 165, duration:
                                       13.680s, episode steps:
                                                                764, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                            -1.000,
                                                                                                                                                    0.000], mean action:
144768/500000; episode: 166, duration: 14.555s, episode steps: 820, steps per second: 56, episode reward: -21.000, mean reward: -0.026
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
               episode: 167, duration: 13.672s, episode steps:
                                                                764, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                                    0.000], mean action:
146408/500000: episode: 168, duration: 15.514s, episode steps: 876, steps per second: 56, episode reward: -21.000, mean reward: -0.024
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
                                                                792, steps per second:
148080/500000: episode: 170, duration: 15.577s, episode steps: 880, steps per second: 56, episode reward: -21.000, mean reward: -0.024
                                                                                                                                           [-1.000, 0.000], mean action:
               episode: 171, duration:
                                        15.312s, episode steps:
                                                                871, steps per second:
                                                                                        57, episode reward: -20.000, mean reward: -0.023
                                                                                                                                                    1.000], mean action:
149715/500000:
               episode: 172, duration: 13.657s, episode steps: 764, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
150479/500000:
               episode: 173, duration: 13.871s, episode steps: 764, steps per second: 55, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
151340/500000: episode: 174, duration: 15.272s, episode steps: 861, steps per second: 56, episode reward: -20.000, mean reward: -0.023
                                                                                                                                           [-1.000, 1.000], mean action:
               episode: 175, duration: 13.581s, episode steps:
                                                                764, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                                    0.000], mean action:
152868/500000: episode: 176, duration: 13.721s, episode steps: 764, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
               episode: 177, duration: 15.565s, episode steps: 870, steps per second: 56, episode reward: -20.000, mean reward: -0.023
                                                                                                                                                    1.000], mean action:
154502/500000: episode: 178. duration: 13.597s. episode steps: 764. steps per second: 56. episode reward: -21.000. mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
                                        15.604s, episode steps:
                                                                870, steps per second: 56,
                                                                                            episode reward: -20.000,
                                                                                                                      mean reward: -0.023
                                                                                                                                            -1.000,
156303/500000: episode: 180, duration: 16.653s, episode steps: 931, steps per second: 56, episode reward: -20.000, mean reward: -0.021
                                                                                                                                           [-1.000, 1.000], mean action:
               episode: 181, duration: 18.033s, episode steps: 1018, steps per second: 56, episode reward: -20.000, mean reward: -0.020
158085/500000: episode: 182, duration: 13.566s, episode steps: 764, steps per second: 56, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
                                       14.444s, episode steps:
158905/500000:
               episode: 183, duration:
                                                                820, steps per second: 57,
                                                                                            episode reward: -21.000, mean reward: -0.026
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
159825/500000: episode: 184, duration: 16.251s, episode steps: 920, steps per second: 57, episode reward: -20.000, mean reward: -0.022
                                                                                                                                           [-1.000, 1.000], mean action:
                                       13.537s, episode steps:
                                                                764, steps per second: 56,
                                                                                            episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000,
161353/500000; episode: 186, duration: 13.610s, episode steps: 764, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000, 0.000], mean action:
162177/500000:
               episode: 187, duration:
                                       14.469s, episode steps: 824, steps per second: 57,
                                                                                            episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
162969/500000: episode: 188, duration: 13.827s, episode steps: 792, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
               episode: 189, duration: 14.792s, episode steps: 843, steps per second: 57,
                                                                                            episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000,
164868/500000: episode: 190, duration: 18.672s, episode steps: 1056, steps per second: 57, episode reward: -19.000, mean reward: -0.018 [-1.000, 1.000], mean action:
                                                                842, steps per second: 57, episode reward: -20.000, mean reward: -0.024
               episode: 191, duration:
                                       14.880s, episode steps:
166627/500000:
               episode: 192, duration: 16.113s, episode steps: 917, steps per second: 57, episode reward: -20.000, mean reward: -0.022
                                                                                                                                           [-1.000, 1.000], mean action:
167499/500000:
               episode: 193, duration: 15.678s, episode steps: 872, steps per second: 56, episode reward: -20.000, mean reward: -0.023
                                                                                                                                           [-1.000,
                                                                                                                                                    1.000], mean action:
168319/500000:
               episode: 194, duration: 14.424s, episode steps: 820, steps per second: 57, episode reward: -21.000, mean reward: -0.026
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
                                                                                                                                           [-1.000,
                                                                     steps per second: 56,
                                                                                            episode reward: -21.000, mean reward: -0.022
               episode: 195, duration: 17.183s, episode steps: 970,
                                                                                                                                                    0.000], mean action:
169289/500000:
170100/500000: episode: 196, duration: 14.249s, episode steps: 811, steps per second: 57, episode reward: -21.000, mean reward: -0.026
                                                                                                                                           [-1.000, 0.000], mean action:
               episode: 197, duration: 14.437s, episode steps: 812, steps per second: 56,
                                                                                                                                           [-1.000,
171754/500000: episode: 198, duration: 14.877s, episode steps: 842, steps per second: 57, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000, 1.000], mean action:
172684/500000:
                                       16.692s, episode steps: 930,
                                                                     steps per second: 56,
                                                                                            episode reward: -20.000, mean reward: -0.022
                                                                                                                                            -1.000,
173554/500000:
               episode: 200, duration: 15.602s, episode steps: 870, steps per second: 56, episode reward: -20.000, mean reward: -0.023
                                                                                                                                           [-1.000.
                                                                                                                                                    1.0001, mean action:
174318/500000:
               episode: 201, duration:
                                       13.664s, episode steps: 764,
                                                                     steps per second: 56,
                                                                                            episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
175185/500000: episode: 202, duration: 15.504s, episode steps: 867, steps per second: 56, episode reward: -21.000, mean reward: -0.024
                                                                                                                                           [-1.000, 0.000], mean action:
                                       13.600s,
               episode: 203, duration:
                                                                                            episode reward: -21.000,
175949/500000
                                                episode steps:
                                                                764.
                                                                     steps per second: 56,
                                                                                                                      mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.000], mean action
176713/500000: episode: 204, duration: 13.320s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000, 0.000], mean action:
               episode: 205, duration:
                                       14.539s, episode steps: 825, steps per second:
                                                                                            episode reward: -21.000, mean reward: -0.025
178383/500000:
               episode: 206, duration: 15.031s, episode steps: 845, steps per second: 56, episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000, 0.000], mean action:
179147/500000:
               episode: 207, duration: 13.598s, episode steps: 764, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
179958/500000:
               episode: 208, duration: 14.465s, episode steps: 811, steps per second: 56, episode reward: -21.000, mean reward: -0.026
                                                                                                                                           [-1.000, 0.000], mean action:
               episode: 209, duration: 13.456s, episode steps: 764, steps per second:
181486/500000: episode: 210, duration: 13.556s, episode steps: 764, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
                                                                                                                                                    0.000],
182250/500000
               episode: 211,
                                       13.449s,
                                                episode steps:
                                                                     steps per second:
                                                                                            episode reward: -21.000,
                                                                                                                      mean reward: -0.027
                                                                                                                                             1.000,
183014/500000: episode: 212, duration: 13.764s, episode steps: 764, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                            -1.000,
                                                                                                                                                    0.000], mean action:
                                                                                                                                            -1.000,
                                        13.386s, episode steps:
                                                                     steps per second:
                                                                                            episode reward: -21.000, mean
184570/500000; episode: 214, duration: 13.899s, episode steps: 792, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000, 0.000], mean action:
185363/500000:
               episode: 215, duration: 14.290s, episode steps:
                                                                793, steps per second: 55,
                                                                                            episode reward: -21.000, mean reward: -0.026
                                                                                                                                           [-1.000, 0.000], mean action:
186156/500000: episode: 216, duration: 14.120s, episode steps: 793, steps per second: 56, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 186949/500000: episode: 217, duration: 13.957s, episode steps: 793, steps per second: 57, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action:
```

```
187713/500000: episode: 218, duration: 13.619s, episode steps: 764, steps per second: 56, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
188477/500000: episode: 219, duration: 13.452s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
              episode: 220, duration: 14.715s, episode steps: 843, steps per second: 57, episode reward: -20.000, mean reward: -0.024
190084/500000: episode: 221, duration: 13.536s, episode steps: 764, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                   0.0001, mean action:
190906/500000: episode: 222, duration: 14.482s, episode steps: 822, steps per second: 57, episode reward: -21.000, mean reward: -0.026
                                                                                                                                           [-1.000,
                                                                                                                                                   0.000], mean action:
191670/500000: episode: 223, duration: 13.328s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000,
                                                                                                                                                   0.000], mean action:
              episode: 224, duration: 14.042s, episode steps:
                                                                782, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                                    0.000], mean action:
193284/500000: episode: 225, duration: 14.724s, episode steps: 832, steps per second: 57, episode reward: -21.000, mean reward: -0.025 194242/500000: episode: 226, duration: 16.878s, episode steps: 958, steps per second: 57, episode reward: -20.000, mean reward: -0.021
                                                                                                                                           [-1.000.
                                                                                                                                                   0.0001, mean action:
                                                                                                                                                     .000], mean action:
195091/500000: episode: 227, duration: 14.740s, episode steps: 849, steps per second: 58, episode reward: -21.000, mean reward: -0.025
                                                                                                                                          [-1.000, 0.000], mean action:
              episode: 228, duration: 13.922s, episode steps: 792, steps per second: 57, episode reward: -21.000, mean reward: -
                                                                                                                                            -1.000,
                                                                                                                                                   0.000], mean action:
196675/500000: episode: 229, duration: 13.891s, episode steps: 792, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                   0.0001, mean action:
              episode: 230, duration: 14.497s, episode steps:
                                                                821, steps per second:
                                                                                            episode reward: -21.000, mean reward: -0.026
                                                                                                                                                    0.000], mean action:
198398/500000: episode: 231, duration: 16.002s, episode steps: 902, steps per second: 56, episode reward: -20.000, mean reward: -0.022
                                                                                                                                           [-1.000.
                                                                                                                                                   1.0001, mean action:
              episode: 232, duration:
                                       14.753s, episode steps:
                                                                840, steps per second:
                                                                                        57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                                    0.000], mean action:
200002/500000; episode: 233, duration: 13.313s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                   0.0001, mean action:
              episode: 234, duration: 13.414s, episode steps:
                                                                764, steps per second: 57,
                                                                                            episode reward: -21.000, mean reward: -0.027
                                                                                                                                            -1.000,
                                                                                                                                                    0.000], mean action:
201530/500000: episode: 235, duration: 13.128s, episode steps: 764, steps per second: 58, episode reward: -21.000, mean reward: -0.027
                                                                                                                                          [-1.000,
                                                                                                                                                   0.0001, mean action:
              episode: 236, duration: 13.839s, episode steps:
                                                                                                                                           [-1.000,
202313/500000:
                                                                783, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                                   0.000], mean action:
203125/500000: episode: 237, duration: 14.639s, episode steps: 812, steps per second: 55, episode reward: -21.000, mean reward: -0.026
                                                                                                                                           [-1.000,
                                                                                                                                                   0.000], mean action:
203889/500000: episode: 238, duration: 13.495s, episode steps: 764, steps per second:
                                                                                        57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000,
                                                                                                                                                   0.000], mean action:
204824/500000: episode: 239, duration: 16.482s, episode steps: 935, steps per second: 57, episode reward: -20.000, mean reward: -0.021
                                                                                                                                           [-1.000.
                                                                                                                                                   1.000], mean action:
              episode: 240, duration:
                                       14.816s, episode steps:
                                                                843, steps per second:
                                                                                            episode reward: -21.000,
                                                                                                                     mean reward: -0.025
                                                                                                                                                    0.000], mean action:
206431/500000: episode: 241, duration: 13.511s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                   0.0001, mean action:
207273/500000: episode: 242, duration: 14.725s, episode steps: 842, steps per second: 57, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000,
                                                                                                                                                    1.000], mean action:
208037/500000: episode: 243, duration: 13.440s, episode steps: 764, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                   0.0001, mean action:
               episode: 244, duration: 14.277s, episode steps:
                                                                812, steps per second:
                                                                                            episode reward: -21.000, mean reward: -0.026
209670/500000: episode: 245. duration: 14.761s. episode steps: 821. steps per second: 56. episode reward: -21.000. mean reward: -0.026
                                                                                                                                           [-1.000, 0.000], mean action:
              episode: 246, duration: 13.491s, episode steps: 764, steps per second:
                                                                                            episode reward: -21.000, mean reward: -0.027
                                                                                                                                                    0.000], mean action:
211306/500000: episode: 247, duration: 15.271s, episode steps: 872, steps per second: 57, episode reward: -20.000, mean reward: -0.023
                                                                                                                                           [-1.000.
                                                                                                                                                   1.0001, mean action:
              episode: 248, duration:
                                       13.543s, episode steps: 764, steps per second: 56,
                                                                                            episode reward: -21.000, mean reward: -0.027
                                                                                                                                            -1.000,
                                                                                                                                                    0.000], mean action:
212881/500000: episode: 249, duration: 14.569s, episode steps: 811, steps per second: 56, episode reward: -21.000, mean reward: -0.026
                                                                                                                                           [-1.000.
                                                                                                                                                   0.0001, mean action:
              episode: 250, duration: 14.385s, episode steps: 822, steps per second:
214592/500000: episode: 251, duration: 15.649s, episode steps: 889, steps per second: 57, episode reward: -20.000, mean reward: -0.022
                                                                                                                                          [-1.000, 1.000], mean action:
              episode: 252, duration:
                                       14.381s, episode steps: 820, steps per second:
                                                                                            episode reward: -21.000, mean reward: -0.026
                                                                                                                                                    0.000], mean action:
216264/500000:
              episode: 253, duration: 15.181s, episode steps: 852, steps per second: 56, episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000.
                                                                                                                                                   0.0001, mean action:
217185/500000:
              episode: 254, duration: 16.243s, episode steps: 921, steps per second: 57,
                                                                                            episode reward: -19.000, mean reward: -0.021
                                                                                                                                           [-1.000,
                                                                                                                                                    1.000], mean action:
218024/500000: episode: 255, duration: 14.917s, episode steps: 839, steps per second: 56, episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000.
                                                                                                                                                   0.0001, mean action:
               episode: 256, duration:
                                       13.835s, episode steps:
                                                                                                                     mean reward: -0.027
218807/500000:
                                                               783,
                                                                     steps per second: 57,
                                                                                            episode reward: -21.000,
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
219679/500000: episode: 257, duration: 15.435s, episode steps: 872, steps per second: 56, episode reward: -21.000, mean reward: -0.024
                                                                                                                                           [-1.000,
                                                                                                                                                   0.000], mean action:
              episode: 258, duration: 14.497s, episode steps: 807, steps per second: 56, episode reward: -21.000, mean reward: -0.026
                                                                                                                                                    0.000], mean action:
221250/500000: episode: 259. duration: 13.609s. episode steps: 764. steps per second: 56. episode reward: -21.000. mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                   0.0001, mean action:
              episode: 260, duration:
222161/500000:
                                       16.021s, episode steps: 911, steps per second: 57,
                                                                                            episode reward: -21.000, mean reward: -0.023
                                                                                                                                            -1.000,
                                                                                                                                                    0.000], mean action:
222953/500000; episode: 261, duration: 13.848s, episode steps: 792, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                   0.0001, mean action:
              episode: 262, duration:
                                       15.371s, episode steps: 880, steps per second:
                                                                                            episode reward: -21.000, mean reward: -0.024
                                                                                                                                                    0.000], mean action:
224625/500000: episode: 263, duration: 13.872s, episode steps: 792, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000,
                                                                                                                                                   0.000], mean action:
                                       15.413s, episode steps:
               episode: 264, duration:
                                                                872,
                                                                     steps per second:
                                                                                            episode reward: -20.000,
                                                                                                                     mean reward: -0.023
                                                                                                                                                    1.000], mean action:
226339/500000: episode: 265, duration: 14.916s, episode steps: 842, steps per second: 56, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000,
                                                                                                                                                   1.000], mean action:
              episode: 266, duration:
                                          .568s, episode steps: 981, steps per second: 56, episode reward: -19.000, mean reward: -0.019
                                                                                                                                                    1.000], mean action:
228299/500000; episode: 267, duration: 17.119s, episode steps: 979, steps per second: 57, episode reward: -19.000, mean reward: -0.019
                                                                                                                                           [-1.000.
                                                                                                                                                   1.0001, mean action:
                                       13.430s, episode steps:
229063/500000:
              episode: 268, duration:
                                                                764, steps per second: 57,
                                                                                            episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
229934/500000: episode: 269, duration: 15.280s, episode steps: 871, steps per second: 57, episode reward: -20.000, mean reward: -0.023
                                                                                                                                           [-1.000.
                                                                                                                                                   1.0001, mean action:
              episode: 270, duration: 16.258s, episode steps: 930, steps per second: 57,
                                                                                            episode reward: -20.000, mean reward:
                                                                                                                                                    1.000], mean action:
231811/500000: episode: 271, duration: 16.490s, episode steps: 947, steps per second: 57, episode reward: -20.000, mean reward: -0.021
                                                                                                                                           [-1.000,
                                                                                                                                                   1.000], mean action:
                                       14.939s, episode steps:
               episode: 272, duration:
                                                                843, steps per second: 56,
                                                                                            episode reward: -20.000, mean reward: -0.024
                                                                                                                                                    1.000], mean action:
233496/500000: episode: 273. duration: 14.637s. episode steps: 842. steps per second: 58. episode reward: -20.000. mean reward: -0.024
                                                                                                                                           [-1.000.
                                                                                                                                                   1.0001, mean action:
              episode: 274, duration: 14.951s, episode steps: 848, steps per second:
                                                                                            episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
235136/500000; episode: 275, duration: 14.234s, episode steps: 792, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                   0.0001, mean action:
                                       15.665s, episode steps:
              episode: 276, duration:
                                                                889, steps per second: 57,
                                                                                            episode reward: -20.000, mean reward: -0.022
                                                                                                                                                    1.000], mean action:
236891/500000: episode: 277, duration: 15.092s, episode steps: 866, steps per second: 57, episode reward: -20.000, mean reward: -0.023
                                                                                                                                           [-1.000.
                                                                                                                                                   1.0001, mean action:
              episode: 278, duration:
                                       14.319s, episode steps: 812, steps per second:
238661/500000: episode: 279, duration: 17.006s, episode steps: 958, steps per second: 56, episode reward: -20.000, mean reward: -0.021
                                                                                                                                          [-1.000, 1.000], mean action:
              episode: 280, duration:
                                        15.674s, episode steps:
                                                                898, steps per second:
                                                                                        57, episode reward: -21.000, mean reward: -0.023
                                                                                                                                                    0.000], mean action:
240361/500000; episode: 281, duration: 14.120s, episode steps: 802, steps per second: 57, episode reward: -21.000, mean reward: -0.026
                                                                                                                                           [-1.000.
                                                                                                                                                   0.0001, mean action:
              episode: 282, duration: 18.977s, episode steps: 1083, steps per second: 57, episode reward: -18.000, mean reward: -0.017
                                                                                                                                           [-1.000,
                                                                                                                                                     1.000], mean action:
242358/500000: episode: 283, duration: 16.083s, episode steps: 914, steps per second: 57, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action:
              episode: 284, duration:
                                       17.232s, episode steps: 972, steps per second: 56, episode reward: -19.000, mean reward: -0.020
                                                                                                                                                    1.000], mean action:
244208/500000: episode: 285, duration: 15.596s, episode steps: 878, steps per second: 56, episode reward: -21.000, mean reward: -0.024
                                                                                                                                           [-1.000,
                                                                                                                                                   0.000], mean action:
              episode: 286, duration: 16.311s, episode steps: 920, steps per second: 56, episode reward: -19.000, mean reward: -0.021
                                                                                                                                                    1.000], mean action:
246054/500000: episode: 287. duration: 16.180s. episode steps: 926. steps per second: 57. episode reward: -20.000. mean reward: -0.022
                                                                                                                                           [-1.000.
                                                                                                                                                   1.0001, mean action:
                                       20.718s, episode steps:
                                                                                         56, episode reward: -18.000, mean reward: -0.015
               episode: 288, duration:
                                                                1166, steps per second:
                                                                                                                                            [-1.000,
248170/500000; episode: 289, duration: 16.668s, episode steps: 950, steps per second: 57, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action:
              episode: 290, duration:
                                       15.646s, episode steps: 901, steps per second: 58, episode reward: -20.000, mean reward: -0.022
249941/500000: episode: 291, duration: 15.244s, episode steps: 870, steps per second: 57, episode reward: -20.000, mean reward: -0.023
                                                                                                                                          [-1.000,
                                                                                                                                                   1.000], mean action:
                                       15.714s, episode steps:
250841/500000:
               episode: 292, duration:
                                                               900, steps per second: 57,
                                                                                            episode reward: -20.000, mean reward: -0.022
                                                                                                                                                    1.000], mean action:
251731/500000: episode: 293, duration: 15.739s, episode steps: 890, steps per second: 57, episode reward: -20.000, mean reward: -0.022
                                                                                                                                          [-1.000,
                                                                                                                                                   1.000], mean action:
                                       16.300s, episode steps: 926, steps per second: 57,
                                                                                            episode reward: -20.000, mean reward: -0.022
253674/500000; episode: 295, duration: 17.579s, episode steps: 1017, steps per second: 58, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action:
              episode: 296, duration: 14.146s, episode steps: 811, steps per second: 57, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action:
254485/500000:
255357/500000: episode: 297, duration: 15.448s, episode steps: 872, steps per second: 56, episode reward: -21.000, mean reward: -0.024
                                                                                                                                          [-1.000, 0.000], mean action:
                                                                                            episode reward: -21.000, mean reward: -0.027
               episode: 298, duration: 13.482s, episode steps:
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
256913/500000: episode: 299, duration: 14.083s, episode steps: 792, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                          [-1.000, 0.000], mean action:
              episode: 300, duration:
                                       16.997s, episode steps:
                                                                967, steps per second:
                                                                                        57, episode reward: -19.000, mean reward: -0.020
258644/500000; episode: 301, duration: 13.661s, episode steps: 764, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                   0.0001, mean action:
259556/500000:
              episode: 302, duration: 16.047s, episode steps: 912, steps per second: 57, episode reward: -21.000, mean reward: -0.023
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
260412/500000: episode: 303, duration: 15.270s, episode steps: 856, steps per second: 56, episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000.
                                                                                                                                                   0.0001, mean action:
                                                                                                                                           r-1.000,
               episode: 304, duration: 14.562s, episode steps: 822, steps per second: 56, episode reward: -21.000, mean reward: -0.026
                                                                                                                                                   0.000], mean action:
261234/500000:
262126/500000: episode: 305, duration: 15.998s, episode steps: 892, steps per second: 56, episode reward: -20.000, mean reward: -0.022
                                                                                                                                           [-1.000, 1.000], mean action:
              episode: 306, duration: 15.131s, episode steps: 843, steps per second: 56, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000,
263811/500000: episode: 307, duration: 15.146s, episode steps: 842, steps per second: 56, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000, 1.000], mean action:
264681/500000:
                                       15.516s, episode steps: 870,
                                                                     steps per second: 56,
                                                                                            episode reward: -20.000, mean reward: -0.023
                                                                                                                                           -1.000,
265501/500000.
              episode: 309, duration: 14.736s, episode steps: 820, steps per second: 56, episode reward: -21.000, mean reward: -0.026
                                                                                                                                           [-1.000.
                                                                                                                                                   0.0001, mean action:
266455/500000:
              episode: 310, duration: 16.703s, episode steps: 954, steps per second: 57,
                                                                                                                                           [-1.000,
                                                                                                                                                    1.000], mean action:
267331/500000: episode: 311, duration: 15.593s, episode steps: 876, steps per second: 56, episode reward: -21.000, mean reward: -0.024
                                                                                                                                           [-1.000,
                                                                                                                                                   0.000], mean action:
                                                                     steps per second: 56, episode reward: -21.000,
268123/500000
               episode: 312, duration: 14.187s, episode steps:
                                                                792.
                                                                                                                     mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.000], mean action
268933/500000: episode: 313, duration: 14.481s, episode steps: 810, steps per second: 56, episode reward: -21.000, mean reward: -0.026
                                                                                                                                           [-1.000, 0.000], mean action:
               episode: 314, duration: 15.507s, episode steps: 871, steps per second: 56, episode reward: -20.000, mean reward: -0.023
270646/500000:
              episode: 315, duration: 15.013s, episode steps: 842, steps per second: 56, episode reward: -20.000, mean reward: -0.024
                                                                                                                                          [-1.000, 1.000], mean action:
271670/500000:
               episode: 316, duration: 18.323s, episode steps: 1024, steps per second: 56, episode reward: -19.000, mean reward: -0.019
                                                                                                                                           [-1.000,
                                                                                                                                                     1.000], mean action
272462/500000:
              episode: 317, duration: 14.059s, episode steps: 792, steps per second: 56, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
              episode: 318, duration: 16.726s, episode steps: 927, steps per second: 55, episode reward: -21.000, mean reward: -0.023
                                                                                                                                           [-1.000,
274307/500000: episode: 319, duration: 16.257s, episode steps: 918, steps per second: 56, episode reward: -20.000, mean reward: -0.022
                                                                                                                                           [-1.000, 1.000], mean action:
                                                                                                                                                    0.000],
275127/500000:
               episode: 320, duration:
                                       14.561s, episode steps:
                                                                820,
                                                                     steps per second: 56,
                                                                                            episode reward: -21.000,
                                                                                                                     mean reward: -0.026
                                                                                                                                           -1.000,
275965/500000: episode: 321, duration: 14.916s, episode steps: 838, steps per second: 56, episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000, 0.000], mean action:
                                                                                                                                           -1.000,
               episode: 322, duration: 14.087s, episode steps:
                                                                792, steps per second: 56, episode reward: -21.000, mean reward: -0.027
277660/500000; episode: 323, duration: 16.042s, episode steps: 903, steps per second: 56, episode reward: -21.000, mean reward: -0.023
                                                                                                                                          [-1.000, 0.000], mean action:
278627/500000:
              episode: 324, duration: 17.254s, episode steps: 967, steps per second: 56, episode reward: -19.000, mean reward: -0.020
                                                                                                                                           [-1.000, 1.000], mean action:
279497/500000: episode: 325, duration: 15.368s, episode steps: 870, steps per second: 57, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 280317/500000: episode: 326, duration: 14.549s, episode steps: 820, steps per second: 56, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action:
```

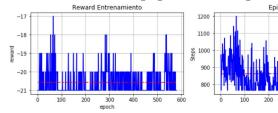
```
281221/500000: episode: 327, duration: 15.745s, episode steps: 904, steps per second: 57, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action:
282065/500000: episode: 328, duration: 14.764s, episode steps: 844, steps per second: 57, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action:
              episode: 329, duration: 17.056s, episode steps: 963, steps per second: 56, episode reward: -21.000, mean reward: -0.022
283870/500000: episode: 330, duration: 14.897s, episode steps: 842, steps per second: 57, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000,
                                                                                                                                                    1.0001, mean action:
284713/500000: episode: 331, duration: 14.862s, episode steps: 843, steps per second: 57, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000,
                                                                                                                                                    1.000], mean action:
285555/500000: episode: 332, duration: 14.994s, episode steps: 842, steps per second: 56, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000, 1.000], mean action:
              episode: 333, duration:
                                       17.811s, episode steps: 1005, steps per second: 56, episode reward: -20.000, mean reward: -0.020
287464/500000: episode: 334, duration: 16.025s, episode steps: 904, steps per second: 56, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action: 288308/500000: episode: 335, duration: 14.863s, episode steps: 844, steps per second: 57, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action:
289210/500000: episode: 336, duration: 15.915s, episode steps: 902, steps per second: 57, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action:
              episode: 337, duration: 14.557s, episode steps:
                                                                820, steps per second: 56, episode reward: -21.000, mean reward: -0.026
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
290920/500000: episode: 338, duration: 15.997s, episode steps: 890, steps per second: 56, episode reward: -20.000, mean reward: -0.022
                                                                                                                                           [-1.000.
                                                                                                                                                    1.0001, mean action:
              episode: 339, duration: 16.020s, episode steps: 898, steps per second: 56, episode reward: -20.000, mean reward: -0.022
                                                                                                                                                    1.000], mean action:
292723/500000: episode: 340, duration: 16.059s, episode steps: 905, steps per second: 56, episode reward: -21.000, mean reward: -0.023
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
              episode: 341, duration:
                                       14.075s, episode steps:
                                                                792, steps per second:
                                                                                        56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                                    0.000], mean action:
294341/500000; episode: 342, duration: 14.663s, episode steps: 826, steps per second: 56, episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
              episode: 343, duration: 13.978s, episode steps:
                                                                792, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                             -1.000,
                                                                                                                                                    0.000], mean action:
296059/500000: episode: 344, duration: 16.529s, episode steps: 926, steps per second: 56, episode reward: -20.000, mean reward: -0.022
                                                                                                                                           [-1.000,
                                                                                                                                                    1.0001, mean action:
                                                                                                                                           [-1.000,
              episode: 345, duration: 14.889s, episode steps: 839, steps per second: 56, episode reward: -21.000, mean reward: -0.025
                                                                                                                                                    0.000], mean action:
297709/500000: episode: 346, duration: 14.500s, episode steps: 811, steps per second: 56, episode reward: -21.000, mean reward: -0.026
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
              episode: 347, duration: 14.887s, episode steps: 842, steps per second: 57, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000,
                                                                                                                                                    1.000], mean action:
299394/500000: episode: 348, duration: 14.623s, episode steps: 843, steps per second: 58, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000.
                                                                                                                                                    1.000], mean action:
              episode:
                        349, duration:
                                        14.583s, episode steps:
                                                                822, steps per second: 56, episode reward: -21.000, mean reward: -0.026
                                                                                                                                                     .000], mean action:
301008/500000: episode: 350, duration: 14.410s, episode steps: 792, steps per second: 55, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
              episode: 351, duration: 15.036s, episode steps: 844, steps per second: 56, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000,
                                                                                                                                                    1.000], mean action:
302644/500000: episode: 352, duration: 14.215s, episode steps: 792, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
               episode: 353, duration: 18.001s, episode steps:
                                                                1025, steps per second: 57, episode reward: -20.000, mean reward: -0.020
304509/500000: episode: 354, duration: 14.873s, episode steps: 840, steps per second: 56, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
              episode: 355, duration: 15.551s, episode steps: 889, steps per second: 57, episode reward: -20.000, mean reward: -0.022
                                                                                                                                                    1.000], mean action:
306269/500000: episode: 356, duration: 15.126s, episode steps: 871, steps per second: 58, episode reward: -20.000, mean reward: -0.023
                                                                                                                                           [-1.000.
                                                                                                                                                    1.0001, mean action:
              episode: 357, duration: 16.394s, episode steps: 931, steps per second: 57, episode reward: -20.000, mean reward: -0.021
                                                                                                                                                    1.000], mean action:
308043/500000: episode: 358, duration: 15.098s, episode steps: 843, steps per second: 56, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000.
                                                                                                                                                    1.0001, mean action:
              episode: 359, duration: 16.274s, episode steps: 927, steps per second: 57,
309753/500000: episode: 360, duration: 13.707s, episode steps: 783, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000, 0.000], mean action:
              episode: 361, duration:
                                       18.317s, episode steps: 1039, steps per second: 57, episode reward: -20.000, mean reward:
                                                                                                                                                     1.000], mean action:
311755/500000; episode: 362, duration: 16.957s, episode steps: 963, steps per second: 57, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action:
312566/500000:
              episode: 363, duration: 14.099s, episode steps: 811, steps per second: 58, episode reward: -21.000, mean reward: -0.026
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
313453/500000: episode: 364, duration: 15.703s, episode steps: 887, steps per second: 56, episode reward: -21.000, mean reward: -0.024
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
               episode: 365, duration: 14.692s, episode steps: 842, steps per second: 57,
                                                                                            episode reward: -20.000, mean reward: -0.024
314295/500000.
                                                                                                                                           [-1.000,
                                                                                                                                                    1.000], mean action:
315137/500000: episode: 366, duration: 14.871s, episode steps: 842, steps per second: 57, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000,
                                                                                                                                                    1.000], mean action:
              episode: 367, duration: 13.953s, episode steps: 783, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                                    0.000], mean action:
316788/500000: episode: 368, duration: 15.336s, episode steps: 868, steps per second: 57, episode reward: -21.000, mean reward: -0.024
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
317654/500000:
              episode: 369, duration:
                                       15.192s, episode steps: 866, steps per second:
                                                                                            episode reward: -20.000, mean reward: -0.023
                                                                                                                                            -1.000,
318498/500000; episode: 370, duration: 14.906s, episode steps: 844, steps per second: 57, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000.
                                                                                                                                                    1.0001, mean action:
              episode: 371, duration: 13.911s, episode steps:
                                                                792, steps per second: 57,
                                                                                            episode reward: -21.000, mean reward: -0.027
                                                                                                                                                    0.000], mean action:
320218/500000: episode: 372, duration: 16.085s, episode steps: 928, steps per second: 58, episode reward: -21.000, mean reward: -0.023
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
                        373, duration:
                                       14.754s, episode steps:
                                                                     steps per second:
                                                                                            episode reward: -20.000,
                                                                                                                      mean reward: -0.024
                                                                                                                                            -1.000,
                                                                                                                                                    1.000], mean action:
321964/500000: episode: 374, duration: 15.957s, episode steps: 903, steps per second: 57, episode reward: -21.000, mean reward: -0.023
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
              episode: 375, duration: 16.114s, episode steps: 919, steps per second:
                                                                                            episode reward: -20.000, mean reward: -0.022
                                                                                                                                                    1.000], mean action:
323815/500000: episode: 376, duration: 16.221s, episode steps: 932, steps per second: 57,
                                                                                            episode reward: -20.000, mean reward: -0.021
                                                                                                                                           [-1.000.
                                                                                                                                                    1.0001, mean action:
324738/500000:
              episode: 377, duration:
                                       16.331s, episode steps: 923, steps per second: 57,
                                                                                            episode reward: -21.000, mean reward: -0.023
                                                                                                                                            [-1.000,
                                                                                                                                                    0.000], mean action:
325576/500000: episode: 378, duration: 15.150s, episode steps: 838, steps per second: 55, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000.
                                                                                                                                                    1.0001, mean action:
              episode: 379, duration: 15.937s, episode steps: 896, steps per second: 56,
                                                                                            episode reward: -21.000, mean reward:
                                                                                                                                             -1.000,
                                                                                                                                                    0.000], mean action:
327376/500000: episode: 380, duration: 15.971s, episode steps: 904, steps per second: 57, episode reward: -21.000, mean reward: -0.023
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
                                       15.514s, episode steps:
                                                                     steps per second:
               episode: 381, duration:
                                                                                            episode reward: -20.000,
                                                                                                                      mean reward: -0.022
                                                                                                                                                    1.000], mean action:
329197/500000:
              episode: 382. duration: 16.233s. episode steps: 932. steps per second: 57.
                                                                                            episode reward: -20.000, mean reward: -0.021
                                                                                                                                           [-1.000.
                                                                                                                                                    1.0001, mean action:
              episode: 383, duration: 13.657s, episode steps:
                                                                783, steps per second:
                                                                                            episode reward: -21.000, mean reward: -0.027
                                                                                                                                            -1.000,
                                                                                                                                                    0.000], mean action:
330806/500000: episode: 384, duration: 14.564s, episode steps: 826, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
              episode: 385, duration: 13.806s, episode steps:
                                                                792, steps per second:
                                                                                            episode reward: -21.000, mean reward: -0.027
                                                                                                                                                    0.000], mean action:
                                                                                            episode reward: -20.000, mean reward: -0.022
332521/500000: episode: 386, duration: 16.144s, episode steps: 923, steps per second: 57,
                                                                                                                                           [-1.000.
                                                                                                                                                    1.0001, mean action:
                                                                792, steps per second:
334203/500000: episode: 388, duration: 15.782s, episode steps: 890, steps per second: 56, episode reward: -20.000, mean reward: -0.022
                                                                                                                                           [-1.000, 1.000], mean action:
              episode: 389, duration:
                                        13.875s, episode steps:
                                                                783, steps per second:
                                                                                            episode reward: -21.000, mean reward: -0.027
                                                                                                                                                    0.000], mean action:
335934/500000:
              episode: 390, duration: 16.679s, episode steps: 948, steps per second: 57,
                                                                                            episode reward: -19.000, mean reward: -0.020
                                                                                                                                           [-1.000.
                                                                                                                                                    1.0001, mean action:
              episode: 391, duration: 14.494s, episode steps: 826, steps per second:
                                                                                            episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
337584/500000: episode: 392, duration: 14.585s, episode steps: 824, steps per second: 56, episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
              episode: 393, duration: 14.535s, episode steps: 826, steps per second: 57,
                                                                                            episode reward: -21.000, mean reward: -0.025
                                                                                                                                                    0.000], mean action:
339193/500000: episode: 394, duration: 13.597s, episode steps: 783, steps per second: 58, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
              episode: 395, duration: 14.071s, episode steps:
                                                                792, steps per second:
                                                                                            episode reward: -21.000, mean reward: -0.027
                                                                                                                                                    0.000], mean action:
340811/500000: episode: 396. duration: 14.394s. episode steps: 826. steps per second: 57. episode reward: -21.000. mean reward: -0.025
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
                        397, duration:
                                       14.687s, episode steps:
                                                                     steps per second:
                                                                                            episode reward: -20.000,
                                                                                                                      mean reward: -0.024
                                                                                                                                            -1.000,
342551/500000:
              episode: 398, duration: 15.800s, episode steps: 898, steps per second: 57, episode reward: -20.000, mean reward: -0.022
                                                                                                                                           [-1.000.
                                                                                                                                                    1.0001, mean action:
              episode: 399, duration: 15.022s, episode steps: 825, steps per second: 55,
                                                                                            episode reward: -21.000, mean reward: -0.025
344202/500000: episode: 400, duration: 14.533s, episode steps: 826, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
                                       14.575s, episode steps:
345028/500000:
               episode: 401, duration:
                                                                826,
                                                                     steps per second:
                                                                                            episode reward: -21.000,
                                                                                                                      mean reward: -0.025
                                                                                                                                                    0.000], mean action:
345872/500000: episode: 402, duration: 15.000s, episode steps: 844, steps per second: 56, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000,
                                                                                                                                                    1.000], mean action:
              episode: 403, duration: 15.035s, episode steps: 842, steps per second:
                                                                                            episode reward: -20.000, mean reward: -0.024
                                                                                                                                            -1.000,
347506/500000; episode: 404, duration: 14.297s, episode steps: 792, steps per second: 55, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
348388/500000:
               episode: 405, duration: 15.629s, episode steps: 882, steps per second: 56,
                                                                                            episode reward: -21.000, mean reward: -0.024
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
349180/500000:
              episode: 406, duration: 14.065s, episode steps: 792, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
               episode: 407, duration: 15.571s, episode steps: 889,
                                                                                            episode reward: -20.000,
                                                                                                                                           [-1.000,
350913/500000: episode: 408, duration: 14.606s, episode steps: 844, steps per second: 58, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action:
                                                                842, steps per second:
                                                                                            episode reward: -20.000, mean reward: -0.024
               episode: 409, duration:
                                       14.598s, episode steps:
                                                                                                                                                    1.000], mean action:
352547/500000:
              episode: 410, duration: 14.067s, episode steps: 792, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
353396/500000:
              episode: 411, duration: 14.961s, episode steps: 849, steps per second: 57,
                                                                                            episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
354188/500000:
              episode: 412, duration: 14.101s, episode steps: 792, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000.
                                                                                                                                                    0.0001, mean action:
                                                                                                                                           [-1.000,
                                                                                            episode reward: -21.000,
               episode: 413, duration: 14.512s, episode steps: 826,
                                                                                                                                                    0.000], mean action:
355014/500000:
                                                                                                                      mean reward: -0.025
355856/500000: episode: 414, duration: 14.570s, episode steps: 842, steps per second: 58, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000, 1.000], mean action:
               episode: 415, duration: 14.756s, episode steps: 842,
                                                                     steps per second: 57,
                                                                                            episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000,
357481/500000: episode: 416, duration: 13.831s, episode steps: 783, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000, 0.000], mean action:
358273/500000:
               episode: 417, duration:
                                       13.989s, episode steps:
                                                                     steps per second:
                                                                                            episode reward: -21.000, mean reward: -0.027
                                                                                                                                            -1.000,
                                                                                                                                                    0.000], mean action:
359065/500000.
              episode: 418, duration: 13.801s, episode steps: 792, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000, 0.000], mean action:
359886/500000:
              episode: 419, duration: 14.372s, episode steps: 821, steps per second: 57,
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
360747/500000:
              episode: 420, duration: 15.224s, episode steps: 861, steps per second: 57, episode reward: -20.000, mean reward: -0.023
                                                                                                                                           [-1.000, 1.000], mean action:
               episode: 421, duration: 14.558s, episode steps: 838,
                                                                                            episode reward: -20.000,
361585/500000
                                                                     steps per second: 58,
                                                                                                                      mean reward: -0.024
                                                                                                                                           [-1.000.
                                                                                                                                                    1.000], mean action
362411/500000: episode: 422, duration: 14.257s, episode steps: 826, steps per second: 58, episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000, 0.000], mean action:
               episode: 423, duration: 18.041s, episode steps: 1036, steps per second: 57, episode reward: -20.000, mean reward: -
              episode: 424, duration: 14.451s, episode steps: 824, steps per second: 57, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
364271/500000:
365097/500000:
               episode: 425, duration: 14.742s, episode steps: 826, steps per second: 56,
                                                                                            episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
365950/500000:
              episode: 426, duration: 14.825s, episode steps: 853, steps per second: 58, episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000, 0.000], mean action:
              episode: 427, duration: 14.543s, episode steps: 839, steps per second: 58, episode reward: -21.000, mean reward: -0.025
367615/500000: episode: 428, duration: 14.465s, episode steps: 826, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000, 0.000], mean action:
                                                                                                                                            -1.000,
                                                                                                                                                    0.000],
368439/500000:
               episode: 429,
                                       14.186s, episode steps:
                                                                824.
                                                                     steps per second:
                                                                                            episode reward: -21.000,
                                                                                                                      mean reward: -0.025
369283/500000:
              episode: 430, duration: 14.690s, episode steps: 844, steps per second: 57, episode reward: -20.000, mean reward: -0.024
                                                                                                                                           [-1.000,
                                                                                                                                                    1.000], mean action:
                                                                                                                                            -1.000,
               episode: 431, duration: 14.775s, episode steps: 844, steps per second:
                                                                                            episode reward: -20.000, mean reward: -0.024
370953/500000; episode: 432, duration: 14.404s, episode steps: 826, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                           [-1.000, 0.000], mean action:
371745/500000:
              episode: 433, duration: 14.049s, episode steps:
                                                                792, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                           [-1.000,
                                                                                                                                                    0.000], mean action:
372626/500000: episode: 434, duration: 15.389s, episode steps: 881, steps per second: 57, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 373470/500000: episode: 435, duration: 14.753s, episode steps: 844, steps per second: 57, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action:
```

```
374262/500000: episode: 436, duration: 13.889s, episode steps: 792, steps per second: 57, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
375054/500000: episode: 437, duration: 13.797s, episode steps: 792, steps per second: 57, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
              episode: 438, duration: 13.584s, episode steps:
                                                               784, steps per second: 58, episode reward: -21.000, mean reward: -0.027
376664/500000: episode: 439, duration: 14.539s, episode steps: 826, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000, 0.000], mean action:
              episode: 440, duration: 14.579s, episode steps: 824, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000,
                                                                                                                                                  0.000], mean action:
378341/500000: episode: 441, duration: 15.144s, episode steps: 853, steps per second: 56, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000,
                                                                                                                                                  0.000], mean action:
              episode: 442, duration: 15.064s, episode steps:
                                                                843, steps per second: 56, episode reward: -21.000, mean reward: -0.025
                                                                                                                                                  0.000], mean action:
379976/500000: episode: 443. duration: 14.072s. episode steps: 792. steps per second: 56. episode reward: -21.000. mean reward: -0.027
                                                                                                                                          [-1.000.
                                                                                                                                                  0.0001, mean action:
              episode: 444, duration: 13.934s, episode steps:
                                                                792, steps per second:
                                                                                       57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                                  0.000], mean action:
381594/500000: episode: 445, duration: 14.689s, episode steps: 826, steps per second: 56, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000, 0.000], mean action:
              episode: 446, duration: 14.125s, episode steps:
                                                               792, steps per second: 56, episode reward: -21.000, mean reward:
                                                                                                                                           -1.000,
                                                                                                                                                  0.000], mean action:
383179/500000: episode: 447, duration: 13.849s, episode steps: 793, steps per second: 57, episode reward: -21.000, mean reward: -0.026
                                                                                                                                         [-1.000.
                                                                                                                                                  0.0001, mean action:
              episode: 448, duration: 13.738s, episode steps:
                                                               783, steps per second:
                                                                                           episode reward: -21.000, mean reward: -0.027
                                                                                                                                                  0.000], mean action:
384788/500000: episode: 449, duration: 14.551s, episode steps: 826, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000.
                                                                                                                                                  0.0001, mean action:
              episode: 450, duration:
                                       14.591s, episode steps:
                                                                825, steps per second:
                                                                                       57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                                  0.000], mean action:
386438/500000:
              episode: 451, duration: 14.406s, episode steps: 825, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000.
                                                                                                                                                  0.0001, mean action:
              episode: 452, duration: 14.497s, episode steps:
                                                               824, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                           -1.000,
                                                                                                                                                  0.000], mean action:
388087/500000: episode: 453, duration: 14.553s, episode steps: 825, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000.
                                                                                                                                                  0.0001, mean action:
                                                                                                                                         [-1.000,
              episode: 454, duration: 14.582s, episode steps: 826, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                                  0.000], mean action:
389705/500000: episode: 455, duration: 13.854s, episode steps: 792, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                          [-1.000,
                                                                                                                                                  0.000], mean action:
              episode: 456, duration: 14.508s, episode steps: 826, steps per second:
                                                                                       57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000,
                                                                                                                                                  0.000], mean action:
391429/500000: episode: 457, duration: 16.009s, episode steps: 898, steps per second: 56, episode reward: -20.000, mean reward: -0.022
                                                                                                                                         [-1.000.
                                                                                                                                                  1.000], mean action:
                                       14.902s, episode steps:
              episode:
                       458, duration:
                                                               844, steps per second:
                                                                                           episode reward: -20.000,
                                                                                                                    mean reward: -0.024
                                                                                                                                                    .000], mean action:
393122/500000: episode: 459, duration: 14.905s, episode steps: 849, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000.
                                                                                                                                                  0.0001, mean action:
              episode: 460, duration: 14.077s, episode steps: 792, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                         [-1.000,
                                                                                                                                                  0.000], mean action:
394739/500000: episode: 461, duration: 14.540s, episode steps: 825,
                                                                    steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000.
                                                                                                                                                  0.0001, mean action:
                                       15.027s, episode steps:
               episode: 462, duration:
                                                               843, steps per second:
                                                                                           episode reward: -20.000, mean reward:
396407/500000: episode: 463. duration: 14.833s. episode steps: 825. steps per second: 56. episode reward: -21.000. mean reward: -0.025
                                                                                                                                         [-1.000.
                                                                                                                                                  0.0001, mean action:
              episode: 464, duration: 13.926s, episode steps:
                                                                                           episode reward: -21.000, mean reward: -0.027
                                                                                                                                                  0.000], mean action:
                                                                792, steps per second:
398025/500000: episode: 465, duration: 14.551s, episode steps: 826, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000.
                                                                                                                                                  0.0001, mean action:
              episode: 466, duration: 14.586s, episode steps: 826, steps per second:
                                                                                           episode reward: -21.000, mean reward: -0.025
                                                                                                                                          -1.000,
                                                                                                                                                  0.000], mean action:
399693/500000: episode: 467, duration: 14.975s, episode steps: 842, steps per second: 56, episode reward: -20.000, mean reward: -0.024
                                                                                                                                         [-1.000.
                                                                                                                                                  1.0001, mean action:
               episode: 468, duration: 14.567s, episode steps: 826, steps per second: 57,
401345/500000: episode: 469, duration: 14.520s, episode steps: 826, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000,
                                                                                                                                                  0.000], mean action:
              episode: 470, duration:
                                       14.500s, episode steps:
                                                               826, steps per second:
                                                                                           episode reward: -21.000, mean reward:
                                                                                                                                                  0.000], mean action:
403030/500000:
              episode: 471, duration: 15.096s, episode steps: 859, steps per second: 57, episode reward: -20.000, mean reward: -0.023
                                                                                                                                         [-1.000.
                                                                                                                                                  1.0001, mean action:
403911/500000:
              episode: 472, duration: 15.663s, episode steps: 881, steps per second: 56, episode reward: -21.000, mean reward: -0.024
                                                                                                                                          [-1.000,
                                                                                                                                                  0.000], mean action:
404737/500000: episode: 473, duration: 14.650s, episode steps: 826, steps per second: 56, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000.
                                                                                                                                                  0.0001, mean action:
               episode: 474, duration:
405590/500000.
                                       15.130s, episode steps: 853,
                                                                     steps per second: 56,
                                                                                           episode reward: -21.000,
                                                                                                                    mean reward: -0.025
                                                                                                                                          [-1.000,
                                                                                                                                                  0.000], mean action
406449/500000: episode: 475, duration: 14.956s, episode steps: 859, steps per second: 57, episode reward: -20.000, mean reward: -0.023
                                                                                                                                         [-1.000,
                                                                                                                                                  1.000], mean action:
              episode: 476, duration: 14.936s, episode steps: 842, steps per second: 56, episode reward: -20.000, mean reward: -0.024
                                                                                                                                                   1.000], mean action:
408117/500000: episode: 477. duration: 14.355s. episode steps: 826. steps per second: 58. episode reward: -21.000. mean reward: -0.025
                                                                                                                                         [-1.000.
                                                                                                                                                  0.0001, mean action:
              episode: 478, duration:
                                       14.816s, episode steps:
                                                                    steps per second: 58,
                                                                                           episode reward: -21.000, mean reward: -0.025
                                                                                                                                          -1.000,
                                                                                                                                                   0.000], mean action:
409814/500000; episode: 479, duration: 14.707s, episode steps: 844, steps per second: 57, episode reward: -20.000, mean reward: -0.024
                                                                                                                                         [-1.000.
                                                                                                                                                  1.0001, mean action:
              episode: 480, duration: 14.730s, episode steps: 843, steps per second:
                                                                                           episode reward: -20.000, mean reward: -0.024
                                                                                                                                                    .000], mean action:
411501/500000: episode: 481, duration: 14.965s, episode steps: 844, steps per second: 56, episode reward: -20.000, mean reward: -0.024
                                                                                                                                         [-1.000,
                                                                                                                                                  1.000], mean action:
               episode: 482, duration:
                                       14.844s, episode steps:
                                                                     steps per second:
                                                                                           episode reward: -20.000,
                                                                                                                    mean reward: -0.024
                                                                                                                                                   1.000], mean action:
413344/500000: episode: 483, duration: 17.559s, episode steps: 999, steps per second: 57, episode reward: -19.000, mean reward: -0.019
                                                                                                                                         [-1.000,
                                                                                                                                                  1.000], mean action:
              episode: 484, duration: 14.266s, episode steps:
                                                               792, steps per second:
                                                                                           episode reward: -21.000, mean reward: -0.027
                                                                                                                                                  0.000], mean action:
414962/500000:
              episode: 485, duration: 14.594s, episode steps: 826, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000.
                                                                                                                                                  0.0001, mean action:
415807/500000:
              episode: 486, duration:
                                       14.927s, episode steps:
                                                               845, steps per second: 57,
                                                                                           episode reward: -21.000, mean reward: -0.025
                                                                                                                                          [-1.000,
                                                                                                                                                  0.000], mean action:
416748/500000:
              episode: 487, duration: 16.795s, episode steps: 941, steps per second: 56, episode reward: -19.000, mean reward: -0.020
                                                                                                                                         [-1.000.
                                                                                                                                                  1.0001, mean action:
              episode: 488, duration: 14.924s, episode steps: 825, steps per second: 55,
                                                                                           episode reward: -21.000,
                                                                                                                                           -1.000,
                                                                                                                                                  0.000], mean action:
                                                                                                                    mean reward:
418356/500000:
              episode: 489, duration: 13.783s, episode steps: 783, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                          [-1.000,
                                                                                                                                                  0.000], mean action:
                                       13.978s, episode steps:
               episode: 490, duration:
                                                                783, steps per second:
                                                                                           episode reward: -21.000,
                                                                                                                    mean reward: -0.027
                                                                                                                                                   0.000], mean action:
419931/500000:
              episode: 491. duration: 14.035s. episode steps: 792. steps per second: 56. episode reward: -21.000. mean reward: -0.027
                                                                                                                                          [-1.000.
                                                                                                                                                  0.0001, mean action:
              episode: 492, duration: 14.515s, episode steps: 826, steps per second:
                                                                                           episode reward: -21.000, mean reward: -0.025
                                                                                                                                          -1.000,
                                                                                                                                                  0.000], mean action:
421601/500000: episode: 493, duration: 14.793s, episode steps: 844, steps per second: 57, episode reward: -20.000, mean reward: -0.024
                                                                                                                                         [-1.000.
                                                                                                                                                  1.0001, mean action:
              episode: 494, duration: 13.844s, episode steps:
                                                               783, steps per second:
                                                                                           episode reward: -21.000, mean reward: -0.027
                                                                                                                                                  0.000], mean action:
423176/500000: episode: 495, duration: 14.042s, episode steps: 792, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                         [-1.000.
                                                                                                                                                  0.0001, mean action:
               episode: 496, duration: 16.089s, episode steps: 906, steps per second: 56,
424908/500000: episode: 497, duration: 14.751s, episode steps: 826, steps per second: 56, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000, 0.000], mean action:
              episode: 498, duration:
                                       18.283s, episode steps:
                                                                                           episode reward: -21.000, mean reward: -0.022
                                                                                                                                                  0.000], mean action:
                                                                    steps per second:
426726/500000:
              episode: 499, duration: 14.924s, episode steps: 844, steps per second: 57, episode reward: -20.000, mean reward: -0.024
                                                                                                                                          [-1.000.
                                                                                                                                                  1.0001, mean action:
              episode: 500, duration: 14.110s, episode steps:
                                                               792, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                          [-1.000,
                                                                                                                                                  0.000], mean action:
428310/500000: episode: 501, duration: 14.229s, episode steps: 792, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                         [-1.000.
                                                                                                                                                  0.0001, mean action:
              episode: 502, duration: 15.104s, episode steps:
                                                                844, steps per second: 56, episode reward: -20.000, mean reward: -0.024
                                                                                                                                                  1.000], mean action:
430032/500000: episode: 503, duration: 15.783s, episode steps: 878, steps per second: 56, episode reward: -21.000, mean reward: -0.024
                                                                                                                                          [-1.000,
                                                                                                                                                  0.000], mean action:
              episode: 504, duration: 16.762s, episode steps: 924, steps per second:
                                                                                       55, episode reward: -21.000, mean reward: -0.023
                                                                                                                                                  0.000], mean action:
431739/500000: episode: 505. duration: 14.204s. episode steps: 783. steps per second: 55. episode reward: -21.000. mean reward: -0.027
                                                                                                                                         [-1.000.
                                                                                                                                                  0.0001, mean action:
                       506, duration:
                                       15.016s, episode steps:
                                                               843, steps per second: 56,
                                                                                           episode reward: -21.000,
                                                                                                                    mean reward: -0.025
                                                                                                                                           -1.000,
433419/500000:
              episode: 507, duration: 14.667s, episode steps: 837, steps per second: 57, episode reward: -20.000, mean reward: -0.024
                                                                                                                                         [-1.000.
                                                                                                                                                  1.0001, mean action:
              episode: 508, duration: 15.049s, episode steps: 844, steps per second: 56,
                                                                                           episode reward: -20.000, mean reward: -0.024
435055/500000: episode: 509, duration: 13.812s, episode steps: 792, steps per second: 57, episode reward: -21.000, mean reward: -0.027
                                                                                                                                         [-1.000,
                                                                                                                                                  0.000], mean action:
                                      14.537s, episode steps:
435879/500000:
               episode: 510, duration:
                                                               824,
                                                                     steps per second:
                                                                                           episode reward: -21.000,
                                                                                                                    mean reward: -0.025
                                                                                                                                                   0.000], mean action:
436793/500000: episode: 511, duration: 16.308s, episode steps: 914, steps per second: 56, episode reward: -21.000, mean reward: -0.023
                                                                                                                                         [-1.000, 0.000], mean action:
              episode: 512, duration: 14.342s, episode steps: 825, steps per second: 58,
                                                                                                                                          [-1.000,
438462/500000; episode: 513, duration: 14.882s, episode steps: 844, steps per second: 57, episode reward: -20.000, mean reward: -0.024
                                                                                                                                         [-1.000, 1.000], mean action:
439287/500000:
              episode: 514, duration: 14.352s, episode steps: 825, steps per second: 57,
                                                                                           episode reward: -21.000, mean reward: -0.025
                                                                                                                                          [-1.000,
                                                                                                                                                  0.000], mean action:
440113/500000: episode: 515, duration: 14.342s, episode steps: 826, steps per second: 58, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000.
                                                                                                                                                  0.0001, mean action:
               episode: 516, duration: 14.243s, episode steps: 821, steps per second: 58,
                                                                                           episode reward: -21.000,
                                                                                                                                          [-1.000,
                                                                                                                                                  0.000], mean action
441760/500000: episode: 517, duration: 14.546s, episode steps: 826, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000, 0.000], mean action:
                                                                                           episode reward: -21.000, mean reward: -0.025
               episode: 518, duration:
                                       14.539s, episode steps:
                                                                                                                                                  0.000], mean action:
                                                                    steps per second:
443410/500000:
              episode: 519, duration: 14.660s, episode steps: 824, steps per second: 56, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000.
                                                                                                                                                  0.0001, mean action:
444202/500000:
              episode: 520, duration: 13.736s, episode steps: 792, steps per second: 58, episode reward: -21.000, mean reward: -0.027
                                                                                                                                          [-1.000,
                                                                                                                                                  0.000], mean action:
445028/500000:
              episode: 521, duration: 14.862s, episode steps: 826, steps per second: 56, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000.
                                                                                                                                                  0.0001, mean action:
                                                                                                                                          r=1.000,
                                                                                           episode reward: -21.000,
445820/500000.
               episode: 522, duration: 14.122s, episode steps: 792,
                                                                           per second: 56,
                                                                                                                                                  0.000], mean action:
                                                                                                                    mean reward: -0.027
446760/500000: episode: 523, duration: 16.399s, episode steps: 940, steps per second: 57, episode reward: -21.000, mean reward: -0.022
                                                                                                                                         [-1.000, 0.000], mean action:
              episode: 524, duration: 14.125s, episode steps:
                                                               792, steps per second: 56,
                                                                                                                                          [-1.000,
448344/500000: episode: 525, duration: 14.044s, episode steps: 792, steps per second: 56, episode reward: -21.000, mean reward: -0.027
                                                                                                                                         [-1.000, 0.000], mean action:
449136/500000:
               episode: 526,
                             duration:
                                       14.450s, episode steps:
                                                                     steps per second: 55,
                                                                                           episode reward: -21.000, mean reward: -0.027
                                                                                                                                          -1.000,
                                                                                                                                                  0.000], mean action:
449989/500000.
              episode: 527, duration: 15.034s, episode steps: 853, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                         [-1.000, 0.000], mean action:
450843/500000:
              episode: 528, duration: 14.952s, episode steps: 854, steps per second:
                                                                                                                                         [-1.000,
                                                                                                                                                  0.000], mean action:
451707/500000:
              episode: 529, duration: 15.303s, episode steps: 864, steps per second: 56, episode reward: -21.000, mean reward: -0.024
                                                                                                                                         [-1.000, 0.000], mean action:
               episode: 530, duration: 15.007s, episode steps:
452551/500000
                                                               844.
                                                                    steps per second: 56,
                                                                                           episode reward: -20.000, mean reward: -0.024
                                                                                                                                         [-1.000.
                                                                                                                                                  1.0001, mean action:
453593/500000: episode: 531, duration: 18.317s, episode steps: 1042, steps per second: 57, episode reward: -19.000, mean reward: -0.018 [-1.000, 1.000], mean action:
              episode: 532, duration: 20.163s, episode steps:
                                                               1137, steps per second: 56, episode reward: -20.000, mean reward: -0.018 [-1.000,
              episode: 533, duration: 15.359s, episode steps: 870, steps per second: 57, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action:
455600/500000:
456503/500000:
               episode: 534, duration: 16.174s, episode steps: 903, steps per second: 56, episode reward: -20.000, mean reward: -0.022 [-1.000,
                                                                                                                                                  1.000], mean action:
457485/500000:
              episode: 535, duration: 17.563s, episode steps: 982, steps per second: 56, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action:
              episode: 536, duration: 20.787s, episode steps: 1178, steps per second: 57, episode reward: -18.000, mean reward: -0.015 [-1.000,
459648/500000: episode: 537, duration: 17.385s, episode steps: 985, steps per second: 57, episode reward: -21.000, mean reward: -0.021 [-1.000, 0.000], mean action:
                                                               1196, steps per second: 56, episode reward: -18.000, mean reward: -0.015 [-1.000,
460844/500000
               episode: 538,
                                       21.250s, episode steps:
461717/500000: episode: 539, duration: 15.304s, episode steps: 873, steps per second: 57, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
              episode: 540, duration: 20.367s, episode steps:
                                                               1159, steps per second: 57, episode reward: -19.000, mean reward: -0.016 [-1.000, 1.000], mean action:
464033/500000: episode: 541, duration: 20.123s, episode steps: 1157, steps per second: 57, episode reward: -20.000, mean reward: -0.017 [-1.000, 1.000], mean action:
465008/500000:
              episode: 542, duration: 16.994s, episode steps: 975, steps per second: 57, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action:
465888/500000: episode: 543, duration: 15.487s, episode steps: 880, steps per second: 57, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 466728/500000: episode: 544, duration: 15.173s, episode steps: 840, steps per second: 55, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
```

```
467748/500000: episode: 545, duration: 17.963s, episode steps: 1020, steps per second: 57, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action:
 468924/500000: episode: 546, duration: 20.822s, episode steps: 1176, steps per second: 56, episode reward: -20.000, mean reward: -0.017 [-1.000, 1.000], mean action: 469841/500000: episode: 547, duration: 16.256s, episode steps: 917, steps per second: 56, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action:
 470891/500000: episode: 548, duration: 18.303s, episode steps: 1050, steps per second: 57, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action:
 471731/500000: episode: 549, duration: 14.721s, episode steps: 840, steps per second: 57, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
 472720/500000: episode: 550, duration: 17.412s, episode steps: 989, steps per second: 57, episode reward: -21.000, mean reward: -0.021 [-1.000, 0.000], mean action:
                  episode: 551, duration: 16.429s, episode steps: 931, steps per second: 57, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action:
 474667/500000: episode: 552, duration: 17.896s, episode steps: 1016, steps per second: 57, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action:
 475613/500000: episode: 553, duration: 16.793s, episode steps: 946, steps per second: 56, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action:
 476480/500000: episode: 554, duration: 15.236s, episode steps: 867, steps per second: 57, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
 477319/500000: episode: 555, duration: 14.839s, episode steps: 839, steps per second: 57, episode reward: -21.000, mean reward: -0.025
                                                                                                                                                              [-1.000, 0.000], mean action:
 478264/500000: episode: 556, duration: 16.708s, episode steps: 945, steps per second: 57, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action: 479090/500000: episode: 557, duration: 14.589s, episode steps: 826, steps per second: 57, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
 48009/500000: episode: 558, duration: 17.471s, episode steps: 1000, steps per second: 57, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action:
 481035/500000: episode: 559, duration: 16.652s, episode steps: 945, steps per second: 57, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action:
 481985/500000; episode: 560, duration: 16.670s, episode steps: 950, steps per second: 57, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action:
 482910/500000: episode: 561, duration: 16.294s, episode steps: 925, steps per second: 57, episode reward: -21.000, mean reward: -0.023
 483839/500000: episode: 562, duration: 16.698s, episode steps: 929, steps per second: 56, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action:
 484811/500000: episode: 563, duration: 16.961s, episode steps: 972, steps per second: 57, episode reward: -21.000, mean reward: -0.022
                                                                                                                                                              [-1.000, 0.000], mean action:
 485763/500000: episode: 564, duration: 16.721s, episode steps: 952, steps per second: 57, episode reward: -21.000, mean reward: -0.022
                                                                                                                                                              [-1.000, 0.000], mean action:
 486686/500000: episode: 565, duration: 16.408s, episode steps: 923, steps per second: 56, episode reward: -20.000, mean reward: -0.022
                                                                                                                                                              [-1.000, 1.000], mean action:
 487647/500000: episode: 566, duration: 17.253s, episode steps: 961, steps per second: 56, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action: 488515/500000: episode: 567, duration: 15.392s, episode steps: 868, steps per second: 56, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
 489359/500000: episode: 568, duration: 15.007s, episode steps: 844, steps per second: 56, episode reward: -20.000, mean reward: -0.024
                                                                                                                                                              [-1.000, 1.000], mean action:
 490152/500000: episode: 569, duration: 14.305s, episode steps: 793, steps per second: 55, episode reward: -21.000, mean reward: -0.026
                                                                                                                                                              [-1.000, 0.000], mean action:
 491169/500000: episode: 570, duration: 18.141s, episode steps: 1017, steps per second: 56, episode reward: -21.000, mean reward: -0.021 [-1.000, 0.000], mean action: 492133/500000: episode: 571, duration: 16.830s, episode steps: 964, steps per second: 57, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action:
 493019/500000: episode: 572, duration: 15.775s, episode steps: 886, steps per second: 56, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
 493961/500000: episode: 573, duration: 16.621s, episode steps: 942, steps per second:
                                                                                                     57, episode reward: -21.000, mean reward: -0.022
 494785/500000: episode: 574, duration: 14.613s, episode steps: 824, steps per second: 56, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
 495717/500000: episode: 575, duration: 16.412s, episode steps: 932, steps per second: 57, episode reward: -20.000, mean reward: -0.021
                                                                                                                                                              [-1.000, 1.000], mean action:
 496597/500000: episode: 576, duration: 15.490s, episode steps: 880, steps per second: 57, episode reward: -21.000, mean reward: -0.024 497464/500000: episode: 577, duration: 15.288s, episode steps: 867, steps per second: 57, episode reward: -21.000, mean reward: -0.024
                                                                                                                                                              [-1.000, 0.000], mean action:
 498307/500000: episode: 578, duration: 14.853s, episode steps: 843, steps per second: 57, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
                  episode: 579, duration: 18.211s, episode steps: 1026, steps per second: 56, episode reward: -21.000, mean reward: -0.020 [-1.000, 0.000], mean action:
done, took 8693.123 seconds
```

300

Pesos guardados en: 202001081340_dqn_PongDeterministic-v0_weights.h5f



Testing for 10 episodes Episode 1: reward: -19.000, steps: 976 Episode 2: reward: -19.000, steps: 1000 Episode 3: reward: -21.000, steps: 1036 Episode 4: reward: -20.000, steps: 1255 Episode 5: reward: -20.000, steps: Episode 6: reward: -20.000, steps: 946 Episode 7: reward: -21.000, steps: 826 Episode 8: reward: -20.000, steps: 974 Episode 9: reward: -21.000, steps: 876 Episode 10: reward: -20.000, steps: 898 mean reward: -20.1 mean_steps: 969.5