```
Training for 2500000 steps ...
     822/2500000: episode: 1, duration: 7.851s, episode steps: 822, steps per second: 105, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action:
    1646/2500000: episode: 2, duration: 2.380s, episode steps: 824, steps per second: 346, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
    2526/2500000: episode: 3, duration: 2.597s, episode steps: 880, steps per second: 339, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
    3447/2500000: episode: 4, duration: 2.656s, episode steps: 921, steps per second: 347, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action:
    4271/2500000: episode: 5, duration: 2.318s, episode steps: 824, steps per second: 355, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
    5035/2500000: episode: 6, duration: 2.185s, episode steps: 764, steps per second: 350, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
    5859/2500000: episode: 7, duration: 2.296s, episode steps: 824, steps per second: 359, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
    6870/2500000: episode: 8, duration: 2.832s, episode steps: 1011, steps per second: 357, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action:
    7670/2500000: episode: 9, duration: 2.316s, episode steps: 800, steps per second: 345, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
    8462/2500000: episode: 10, duration: 2.265s, episode steps: 792, steps per second: 350, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
    9347/2500000: episode: 11, duration: 2.442s, episode steps: 885, steps per second: 362, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
   10111/2500000: episode: 12, duration: 3.688s, episode steps: 764, steps per second: 207, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
   11022/2500000: episode: 13, duration: 6.215s, episode steps: 911, steps per second: 147, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action:
   11895/2500000: episode: 14, duration: 5.685s, episode steps: 873, steps per second: 154, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
   12816/2500000: episode: 15, duration: 6.183s, episode steps: 921, steps per second: 149, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action:
   13655/2500000: episode: 16, duration: 5.764s, episode steps: 839, steps per second: 146, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
   14708/2500000: episode: 17, duration: 7.056s, episode steps: 1053, steps per second: 149, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action
   15500/2500000: episode: 18, duration: 5.300s, episode steps: 792, steps per second: 149, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
   16636/2500000: episode: 19, duration: 7.697s, episode steps: 1136, steps per second: 148, episode reward: -20.000, mean reward: -0.018 [-1.000, 1.000], mean action
   17428/2500000: episode: 20, duration: 5.245s, episode steps: 792, steps per second: 151, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
   18326/2500000: episode: 21, duration: 5.850s, episode steps: 898, steps per second: 153, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action:
   19212/2500000: episode: 22, duration: 6.045s, episode steps: 886, steps per second: 147, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
   20098/2500000: episode: 23, duration: 6.016s, episode steps: 886, steps per second: 147, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
   21042/2500000: episode: 24, duration: 6.371s, episode steps: 944, steps per second: 148, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action:
   21884/2500000: episode: 25, duration: 5.675s, episode steps: 842, steps per second: 148, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action:
   22792/2500000: episode: 26, duration: 6.097s, episode steps: 908, steps per second: 149, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action:
   23556/2500000: episode: 27, duration: 5.292s, episode steps: 764, steps per second: 144, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
   24349/2500000: episode: 28, duration: 5.478s, episode steps: 793, steps per second: 145, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action:
   25191/2500000: episode: 29, duration: 5.698s, episode steps: 842, steps per second: 148, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action:
   26174/2500000: episode: 30, duration: 6.556s, episode steps: 983, steps per second: 150, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action
   27213/2500000: episode: 31, duration: 6.901s, episode steps: 1039, steps per second: 151, episode reward: -21.000, mean reward: -0.020 [-1.000, 0.000], mean action
   28147/2500000: episode: 32, duration: 6.263s, episode steps: 934, steps per second: 149, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action:
   29169/2500000: episode: 33, duration: 6.893s, episode steps: 1022, steps per second: 148, episode reward: -21.000, mean reward: -0.021 [-1.000, 0.000], mean action
   30012/2500000: episode: 34, duration: 5.669s, episode steps: 843, steps per second: 149, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
   31115/2500000: episode: 35, duration: 7.339s, episode steps: 1103, steps per second: 150, episode reward: -19.000, mean reward: -0.017 [-1.000, 1.000], mean action
   32095/2500000: episode: 36, duration: 6.816s, episode steps: 980, steps per second: 144, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action:
   32921/2500000: episode: 37, duration: 5.593s, episode steps: 826, steps per second: 148, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
   33899/2500000: episode: 38, duration: 6.589s, episode steps: 978, steps per second: 148, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action:
   34992/2500000: episode: 39, duration: 7.381s, episode steps: 1093, steps per second: 148, episode reward: -21.000, mean reward: -0.019 [-1.000, 0.000], mean action
   35872/2500000: episode: 40, duration: 6.036s, episode steps: 880, steps per second: 146, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
   36785/2500000: episode: 41, duration: 6.211s, episode steps: 913, steps per second: 147, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action:
   37671/2500000: episode: 42, duration: 6.070s, episode steps: 886, steps per second: 146, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
   38664/2500000: episode: 43, duration: 6.689s, episode steps: 993, steps per second: 148, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action:
   39485/2500000: episode: 44, duration: 5.743s, episode steps: 821, steps per second: 143, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action:
   40277/2500000: episode: 45, duration: 5.337s, episode steps: 792, steps per second: 148, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
   41088/2500000: episode: 46, duration: 5.734s, episode steps: 811, steps per second: 141, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action:
   42064/2500000: episode: 47, duration: 6.723s, episode steps: 976, steps per second: 145, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action:
   42907/2500000: episode: 48, duration: 5.720s, episode steps: 843, steps per second: 147, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
   43995/2500000: episode: 49, duration: 7.380s, episode steps: 1088, steps per second: 147, episode reward: -20.000, mean reward: -0.018 [-1.000, 1.000], mean action
   44873/2500000: episode: 50, duration: 5.891s, episode steps: 878, steps per second: 149, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action:
   45656/2500000: episode: 51, duration: 5.247s, episode steps: 783, steps per second: 149, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
   46562/2500000: episode: 52, duration: 6.157s, episode steps: 906, steps per second: 147, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action:
   47768/2500000: episode: 53, duration: 8.184s, episode steps: 1206, steps per second: 147, episode reward: -20.000, mean reward: -0.017 [-1.000, 1.000], mean action
   48740/2500000: episode: 54, duration: 6.498s, episode steps: 972, steps per second: 150, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action:
   49877/2500000: episode: 55, duration: 7.710s, episode steps: 1137, steps per second: 147, episode reward: -18.000, mean reward: -0.016 [-1.000, 1.000], mean action
   50984/2500000: episode: 56, duration: 7.419s, episode steps: 1107, steps per second: 149, episode reward: -19.000, mean reward: -0.017 [-1.000, 1.000], mean action
   51808/2500000: episode: 57, duration: 5.469s, episode steps: 824, steps per second: 151, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
   52849/2500000: episode: 58, duration: 7.114s, episode steps: 1041, steps per second: 146, episode reward: -19.000, mean reward: -0.018 [-1.000, 1.000], mean action
   53631/2500000: episode: 59, duration: 5.163s, episode steps: 782, steps per second: 151, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
   54671/2500000: episode: 60, duration: 7.082s, episode steps: 1040, steps per second: 147, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action
   55709/2500000: episode: 61, duration: 6.941s, episode steps: 1038, steps per second: 150, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action
   56732/2500000: episode: 62, duration: 6.816s, episode steps: 1023, steps per second: 150, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action
   57847/2500000: episode: 63, duration: 7.506s, episode steps: 1115, steps per second: 149, episode reward: -19.000, mean reward: -0.017 [-1.000, 1.000], mean action
   58673/2500000: episode: 64, duration: 5.473s, episode steps: 826, steps per second: 151, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
   59578/2500000: episode: 65, duration: 5.929s, episode steps: 905, steps per second: 153, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action:
   60430/2500000: episode: 66, duration: 5.787s, episode steps: 852, steps per second: 147, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
   61300/2500000: episode: 67, duration: 5.859s, episode steps: 870, steps per second: 148, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
   62287/2500000: episode: 68, duration: 6.467s, episode steps: 987, steps per second: 153, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action:
   63145/2500000: episode: 69, duration: 5.925s, episode steps: 858, steps per second: 145, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
   63928/2500000: episode: 70, duration: 5.171s, episode steps: 783, steps per second: 151, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
   64886/2500000: episode: 71, duration: 6.499s, episode steps: 958, steps per second: 147, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action:
   65861/2500000: episode: 72, duration: 6.504s, episode steps: 975, steps per second: 150, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action:
   66703/2500000: episode: 73, duration: 5.538s, episode steps: 842, steps per second: 152, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action:
   67574/2500000: episode: 74, duration: 5.806s, episode steps: 871, steps per second: 150, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
   68529/2500000: episode: 75, duration: 6.410s, episode steps: 955, steps per second: 149, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action:
   69349/2500000: episode: 76, duration: 5.403s, episode steps: 820, steps per second: 152, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action:
   70191/2500000: episode: 77, duration: 5.706s, episode steps: 842, steps per second: 148, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action:
   71201/2500000: episode: 78, duration: 6.858s, episode steps: 1010, steps per second: 147, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action
   72169/2500000: episode: 79, duration: 6.435s, episode steps: 968, steps per second: 150, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action:
   73244/2500000: episode: 80, duration: 7.093s, episode steps: 1075, steps per second: 152, episode reward: -19.000, mean reward: -0.018 [-1.000, 1.000], mean action
   74113/2500000: episode: 81, duration: 5.752s, episode steps: 869, steps per second: 151, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
   75277/2500000: episode: 82, duration: 7.751s, episode steps: 1164, steps per second: 150, episode reward: -20.000, mean reward: -0.017 [-1.000, 1.000], mean action
   76120/2500000: episode: 83, duration: 5.625s, episode steps: 843, steps per second: 150, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action:
   77158/2500000: episode: 84, duration: 7.018s, episode steps: 1038, steps per second: 148, episode reward: -19.000, mean reward: -0.018 [-1.000, 1.000], mean action
   78333/2500000: episode: 85, duration: 7.830s, episode steps: 1175, steps per second: 150, episode reward: -21.000, mean reward: -0.018 [-1.000, 0.000], mean action
   79403/2500000: episode: 86, duration: 7.044s, episode steps: 1070, steps per second: 152, episode reward: -19.000, mean reward: -0.018 [-1.000, 1.000], mean action
   80439/2500000: episode: 87, duration: 6.968s, episode steps: 1036, steps per second: 149, episode reward: -18.000, mean reward: -0.017 [-1.000, 1.000], mean action
   81406/2500000: episode: 88, duration: 6.450s, episode steps: 967, steps per second: 150, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action
   82291/2500000: episode: 89, duration: 5.944s, episode steps: 885, steps per second: 149, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
   83331/2500000: episode: 90, duration: 7.060s, episode steps: 1040, steps per second: 147, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action
   84526/2500000: episode: 91, duration: 8.149s, episode steps: 1195, steps per second: 147, episode reward: -18.000, mean reward: -0.015 [-1.000, 1.000], mean action
   85454/2500000: episode: 92, duration: 6.401s, episode steps: 928, steps per second: 145, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action:
   86854/2500000: episode: 93, duration: 9.326s, episode steps: 1400, steps per second: 150, episode reward: -18.000, mean reward: -0.013 [-1.000, 1.000], mean action
   88141/2500000: episode: 94, duration: 8.696s, episode steps: 1287, steps per second: 148, episode reward: -17.000, mean reward: -0.013 [-1.000, 1.000], mean action
   89587/2500000: episode: 95, duration: 9.593s, episode steps: 1446, steps per second: 151, episode reward: -17.000, mean reward: -0.012 [-1.000, 1.000], mean action
   90536/2500000: episode: 96, duration: 6.378s, episode steps: 949, steps per second: 149, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action:
   91577/2500000: episode: 97, duration: 6.832s, episode steps: 1041, steps per second: 152, episode reward: -19.000, mean reward: -0.018 [-1.000, 1.000], mean action
   92600/2500000: episode: 98, duration: 6.740s, episode steps: 1023, steps per second: 152, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action
   93972/2500000: episode: 99, duration: 9.024s, episode steps: 1372, steps per second: 152, episode reward: -17.000, mean reward: -0.012 [-1.000, 1.000], mean action
   95554/2500000: episode: 100, duration: 10.421s, episode steps: 1582, steps per second: 152, episode reward: -17.000, mean reward: -0.011 [-1.000, 1.000], mean acti
   96946/2500000: episode: 101, duration: 9.262s, episode steps: 1392, steps per second: 150, episode reward: -19.000, mean reward: -0.014 [-1.000, 1.000], mean acti
   97958/2500000: episode: 102, duration: 6.637s, episode steps: 1012, steps per second: 152, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean acti
   99105/2500000: episode: 103, duration: 7.737s, episode steps: 1147, steps per second: 148, episode reward: -20.000, mean reward: -0.017 [-1.000, 1.000], mean acti
  100387/2500000: episode: 104, duration: 8.486s, episode steps: 1282, steps per second: 151, episode reward: -19.000, mean reward: -0.015 [-1.000, 1.000], mean acti
  101954/2500000: episode: 105, duration: 10.328s, episode steps: 1567, steps per second: 152, episode reward: -15.000, mean reward: -0.010 [-1.000, 1.000], mean acti
  103268/2500000: episode: 106, duration: 8.672s, episode steps: 1314, steps per second: 152, episode reward: -19.000, mean reward: -0.014 [-1.000, 1.000], mean acti
  104746/2500000: episode: 107, duration: 9.785s, episode steps: 1478, steps per second: 151, episode reward: -19.000, mean reward: -0.013 [-1.000, 1.000], mean acti
  106557/2500000: episode: 108, duration: 11.726s, episode steps: 1811, steps per second: 154, episode reward: -13.000, mean reward: -0.007 [-1.000, 1.000], mean acti
```

```
108328/2500000: episode: 109, duration: 11.277s, episode steps: 1771, steps per second: 157, episode reward: -14.000, mean reward: -0.008 [-1.000, 1.000], mean acti
110209/2500000: episode: 110, duration: 11.943s, episode steps: 1881, steps per second: 157, episode reward: -14.000, mean reward: -0.007 [-1.000, 1.000], mean acti
111383/2500000: episode: 111, duration: 7.442s, episode steps: 1174, steps per second: 158, episode reward: -18.000, mean reward: -0.015 [-1.000, 1.000], mean actio
113219/2500000: episode: 112, duration: 11.495s, episode steps: 1836, steps per second: 160, episode reward: -13.000, mean reward: -0.015 [-1.000, 1.000], mean acti
114524/2500000: episode: 113, duration: 8.368s, episode steps: 1305, steps per second: 156, episode reward: -19.000, mean reward: -0.015 [-1.000, 1.000], mean actio
116030/2500000: episode: 114, duration: 9.726s, episode steps: 1506, steps per second: 155, episode reward: -17.000, mean reward: -0.011 [-1.000, 1.000], mean actio
117323/2500000: episode: 115, duration: 8.205s, episode steps: 1293, steps per second: 158, episode reward: -20.000, mean reward: -0.015 [-1.000, 1.000], mean actio
119221/2500000: episode: 116, duration: 12.626s, episode steps: 1898, steps per second: 150, episode reward: -15.000, mean reward: -0.008 [-1.000, 1.000], mean acti
120628/2500000: episode: 117, duration: 9.025s, episode steps: 1407, steps per second: 156, episode reward: -18.000, mean reward: -0.013 [-1.000, 1.000], mean actio
122203/2500000: episode: 118, duration: 10.192s, episode steps: 1575, steps per second: 155, episode reward: -20.000, mean reward: -0.013 [-1.000, 1.000], mean acti
123785/2500000: episode: 119, duration: 10.300s, episode steps: 1582, steps per second: 154, episode reward: -16.000, mean reward: -0.010 [-1.000, 1.000], mean acti
125195/2500000: episode: 120, duration: 9.236s, episode steps: 1410, steps per second: 153, episode reward: -20.000, mean reward: -0.014 [-1.000, 1.000], mean actio
126842/2500000: episode: 121, duration: 10.710s, episode steps: 1647, steps per second: 154, episode reward: -16.000, mean reward: -0.010 [-1.000, 1.000], mean acti
128339/2500000: episode: 122, duration: 9.700s, episode steps: 1497, steps per second: 154, episode reward: -16.000, mean reward: -0.011 [-1.000, 1.000], mean actio
129625/2500000: episode: 123, duration: 8.054s, episode steps: 1286, steps per second: 160, episode reward: -19.000, mean reward: -0.015 [-1.000, 1.000], mean actio
131136/2500000: episode: 124, duration: 9.562s, episode steps: 1511, steps per second: 158, episode reward: -18.000, mean reward: -0.012 [-1.000, 1.000], mean actio
132364/2500000: episode: 125, duration: 7.718s, episode steps: 1228, steps per second: 159, episode reward: -19.000, mean reward: -0.015 [-1.000, 1.000], mean actio
133661/2500000: episode: 126, duration: 8.043s, episode steps: 1297, steps per second: 161, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean actio
135576/2500000: episode: 127, duration: 12.339s, episode steps: 1915, steps per second: 155, episode reward: -13.000, mean reward: -0.007 [-1.000, 1.000], mean acti
136657/2500000: episode: 128, duration: 6.811s, episode steps: 1081, steps per second: 159, episode reward: -18.000, mean reward: -0.017 [-1.000, 1.000], mean actio
137712/2500000: episode: 129, duration: 6.688s, episode steps: 1055, steps per second: 158, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean actio
139209/2500000: episode: 130, duration: 9.311s, episode steps: 1497, steps per second: 161, episode reward: -18.000, mean reward: -0.012 [-1.000, 1.000], mean actio
140452/2500000: episode: 131, duration: 7.623s, episode steps: 1243, steps per second: 163, episode reward: -19.000, mean reward: -0.015 [-1.000, 1.000], mean actio
142122/2500000: episode: 132, duration: 10.454s, episode steps: 1670, steps per second: 160, episode reward: -16.000, mean reward: -0.010 [-1.000, 1.000], mean acti
143604/2500000: episode: 133, duration: 9.397s, episode steps: 1482, steps per second: 158, episode reward: -15.000, mean reward: -0.010 [-1.000, 1.000], mean actio
145504/2500000: episode: 134, duration: 11.927s, episode steps: 1900, steps per second: 159, episode reward: -14.000, mean reward: -0.007 [-1.000, 1.000], mean acti
147290/2500000: episode: 135, duration: 11.340s, episode steps: 1786, steps per second: 157, episode reward: -13.000, mean reward: -0.007 [-1.000, 1.000], mean acti
149097/2500000: episode: 136, duration: 11.283s, episode steps: 1807, steps per second: 160, episode reward: -13.000, mean reward: -0.007 [-1.000, 1.000], mean acti
150360/2500000: episode: 137, duration: 8.054s, episode steps: 1263, steps per second: 157, episode reward: -21.000, mean reward: -0.017 [-1.000, 0.000], mean actio
152061/2500000: episode: 138, duration: 10.597s, episode steps: 1701, steps per second: 161, episode reward: -15.000, mean reward: -0.009 [-1.000, 1.000], mean acti
153320/2500000: episode: 139, duration: 7.955s, episode steps: 1259, steps per second: 158, episode reward: -20.000, mean reward: -0.016 [-1.000, 1.000], mean actio
154873/2500000: episode: 140, duration: 9.501s, episode steps: 1553, steps per second: 163, episode reward: -18.000, mean reward: -0.012 [-1.000, 1.000], mean actio
156800/2500000: episode: 141, duration: 12.207s, episode steps: 1927, steps per second: 158, episode reward: -14.000, mean reward: -0.007 [-1.000, 1.000], mean acti
158838/2500000: episode: 142, duration: 12.647s, episode steps: 2038, steps per second: 161, episode reward: -12.000, mean reward: -0.006 [-1.000, 1.000], mean acti
160427/2500000: episode: 143, duration: 9.794s, episode steps: 1589, steps per second: 162, episode reward: -15.000, mean reward: -0.009 [-1.000, 1.000], mean actio
162252/2500000: episode: 144, duration: 11.509s, episode steps: 1825, steps per second: 159, episode reward: -15.000, mean reward: -0.008 [-1.000, 1.000], mean acti
164212/2500000: episode: 145, duration: 12.292s, episode steps: 1960, steps per second: 159, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
165632/2500000: episode: 146, duration: 9.071s, episode steps: 1420, steps per second: 157, episode reward: -18.000, mean reward: -0.013 [-1.000, 1.000], mean actio
167227/2500000: episode: 147, duration: 9.988s, episode steps: 1595, steps per second: 160, episode reward: -19.000, mean reward: -0.012 [-1.000, 1.000], mean actio
169071/2500000: episode: 148, duration: 11.524s, episode steps: 1844, steps per second: 160, episode reward: -14.000, mean reward: -0.008 [-1.000, 1.000], mean acti
170938/2500000: episode: 149, duration: 11.418s, episode steps: 1867, steps per second: 164, episode reward: -14.000, mean reward: -0.007 [-1.000, 1.000], mean acti
172416/2500000: episode: 150, duration: 9.128s, episode steps: 1478, steps per second: 162, episode reward: -18.000, mean reward: -0.012 [-1.000, 1.000], mean actio
174370/2500000: episode: 151, duration: 12.212s, episode steps: 1954, steps per second: 160, episode reward: -13.000, mean reward: -0.007 [-1.000, 1.000], mean acti
176117/2500000: episode: 152, duration: 10.938s, episode steps: 1747, steps per second: 160, episode reward: -15.000, mean reward: -0.009 [-1.000, 1.000], mean acti
178860/2500000: episode: 153, duration: 17.335s, episode steps: 2743, steps per second: 158, episode reward: -7.000, mean reward: -0.003 [-1.000, 1.000], mean actio
180511/2500000: episode: 154, duration: 10.601s, episode steps: 1651, steps per second: 156, episode reward: -19.000, mean reward: -0.012 [-1.000, 1.000], mean acti
182302/2500000: episode: 155, duration: 11.163s, episode steps: 1791, steps per second: 160, episode reward: -19.000, mean reward: -0.011 [-1.000, 1.000], mean acti
184322/2500000: episode: 156, duration: 13.186s, episode steps: 2020, steps per second: 153, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
186510/2500000: episode: 157, duration: 13.699s, episode steps: 2188, steps per second: 160, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
187589/2500000: episode: 158, duration: 6.710s, episode steps: 1079, steps per second: 161, episode reward: -21.000, mean reward: -0.019 [-1.000, 0.000], mean actio
189580/2500000: episode: 159, duration: 12.566s, episode steps: 1991, steps per second: 158, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
191487/2500000: episode: 160, duration: 11.941s, episode steps: 1907, steps per second: 160, episode reward: -20.000, mean reward: -0.010 [-1.000, 1.000], mean acti
193492/2500000: episode: 161, duration: 12.390s, episode steps: 2005, steps per second: 162, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
195598/2500000: episode: 162, duration: 13.115s, episode steps: 2106, steps per second: 161, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
197398/2500000: episode: 163, duration: 11.014s, episode steps: 1800, steps per second: 163, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
198845/2500000: episode: 164, duration: 9.040s, episode steps: 1447, steps per second: 160, episode reward: -21.000, mean reward: -0.015 [-1.000, 0.000], mean actio
200572/2500000: episode: 165, duration: 10.566s, episode steps: 1727, steps per second: 163, episode reward: -19.000, mean reward: -0.011 [-1.000, 1.000], mean acti
202222/2500000: episode: 166, duration: 10.119s, episode steps: 1650, steps per second: 163, episode reward: -16.000, mean reward: -0.010 [-1.000, 1.000], mean acti
204585/2500000: episode: 167, duration: 14.831s, episode steps: 2363, steps per second: 159, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
207036/2500000: episode: 168, duration: 15.270s, episode steps: 2451, steps per second: 161, episode reward: -7.000, mean reward: -0.003 [-1.000, 1.000], mean actio
208572/2500000: episode: 169, duration: 9.676s, episode steps: 1536, steps per second: 159, episode reward: -17.000, mean reward: -0.011 [-1.000, 1.000], mean actio
210358/2500000: episode: 170, duration: 11.085s, episode steps: 1786, steps per second: 161, episode reward: -17.000, mean reward: -0.010 [-1.000, 1.000], mean acti
212081/2500000: episode: 171, duration: 10.425s, episode steps: 1723, steps per second: 165, episode reward: -15.000, mean reward: -0.009 [-1.000, 1.000], mean acti
213832/2500000: episode: 172, duration: 10.922s, episode steps: 1751, steps per second: 160, episode reward: -19.000, mean reward: -0.011 [-1.000, 1.000], mean acti
215523/2500000: episode: 173, duration: 10.470s, episode steps: 1691, steps per second: 162, episode reward: -16.000, mean reward: -0.009 [-1.000, 1.000], mean acti
217570/2500000: episode: 174, duration: 12.589s, episode steps: 2047, steps per second: 163, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
219558/2500000: episode: 175, duration: 12.491s, episode steps: 1988, steps per second: 159, episode reward: -12.000, mean reward: -0.006 [-1.000, 1.000], mean acti
221368/2500000: episode: 176, duration: 11.162s, episode steps: 1810, steps per second: 162, episode reward: -15.000, mean reward: -0.008 [-1.000, 1.000], mean acti
223323/2500000: episode: 177, duration: 12.215s, episode steps: 1955, steps per second: 160, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
225389/2500000: episode: 178, duration: 12.852s, episode steps: 2066, steps per second: 161, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
227155/2500000: episode: 179, duration: 11.053s, episode steps: 1766, steps per second: 160, episode reward: -17.000, mean reward: -0.010 [-1.000, 1.000], mean acti
229921/2500000: episode: 180, duration: 17.270s, episode steps: 2766, steps per second: 160, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean actio
231975/2500000: episode: 181, duration: 13.001s, episode steps: 2054, steps per second: 158, episode reward: -13.000, mean reward: -0.006 [-1.000, 1.000], mean acti
234064/2500000: episode: 182, duration: 13.506s, episode steps: 2089, steps per second: 155, episode reward: -10.000, mean reward: -0.005 [-1.000, 1.000], mean acti
235890/2500000: episode: 183, duration: 11.384s, episode steps: 1826, steps per second: 160, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean acti
237548/2500000: episode: 184, duration: 10.288s, episode steps: 1658, steps per second: 161, episode reward: -15.000, mean reward: -0.009 [-1.000, 1.000], mean acti
239453/2500000: episode: 185, duration: 12.019s, episode steps: 1905, steps per second: 158, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
241514/2500000: episode: 186, duration: 12.721s, episode steps: 2061, steps per second: 162, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
242938/2500000: episode: 187, duration: 9.016s, episode steps: 1424, steps per second: 158, episode reward: -19.000, mean reward: -0.013 [-1.000, 1.000], mean actio
244680/2500000: episode: 188, duration: 10.836s, episode steps: 1742, steps per second: 161, episode reward: -16.000, mean reward: -0.009 [-1.000, 1.000], mean acti
246697/2500000: episode: 189, duration: 12.505s, episode steps: 2017, steps per second: 161, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
248517/2500000: episode: 190, duration: 11.344s, episode steps: 1820, steps per second: 160, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean acti
251055/2500000: episode: 191, duration: 15.804s, episode steps: 2538, steps per second: 161, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
253132/2500000: episode: 192, duration: 13.034s, episode steps: 2077, steps per second: 159, episode reward: -13.000, mean reward: -0.006 [-1.000, 1.000], mean acti
255524/2500000: episode: 193, duration: 14.660s, episode steps: 2392, steps per second: 163, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
258216/2500000: episode: 194, duration: 16.693s, episode steps: 2692, steps per second: 161, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
261213/2500000: episode: 195, duration: 18.512s, episode steps: 2997, steps per second: 162, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean actio
263502/2500000: episode: 196, duration: 14.191s, episode steps: 2289, steps per second: 161, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
266057/2500000: episode: 197, duration: 15.940s, episode steps: 2555, steps per second: 160, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
268113/2500000: episode: 198, duration: 12.604s, episode steps: 2056, steps per second: 163, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
270909/2500000: episode: 199, duration: 17.302s, episode steps: 2796, steps per second: 162, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
272899/2500000: episode: 200, duration: 12.270s, episode steps: 1990, steps per second: 162, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
275128/2500000: episode: 201, duration: 14.172s, episode steps: 2229, steps per second: 157, episode reward: -13.000, mean reward: -0.006 [-1.000, 1.000], mean acti
277069/2500000: episode: 202, duration: 12.053s, episode steps: 1941, steps per second: 161, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
279078/2500000: episode: 203, duration: 12.390s, episode steps: 2009, steps per second: 162, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
281014/2500000: episode: 204, duration: 12.098s, episode steps: 1936, steps per second: 160, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean acti
283922/2500000: episode: 205, duration: 18.278s, episode steps: 2908, steps per second: 159, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
286036/2500000: episode: 206, duration: 13.066s, episode steps: 2114, steps per second: 162, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
288522/2500000: episode: 207, duration: 15.162s, episode steps: 2486, steps per second: 164, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
291350/2500000: episode: 208, duration: 17.470s, episode steps: 2828, steps per second: 162, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean acti
294114/2500000: episode: 209, duration: 17.004s, episode steps: 2764, steps per second: 163, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
296817/2500000: episode: 210, duration: 16.733s, episode steps: 2703, steps per second: 162, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
299187/2500000: episode: 211, duration: 14.635s, episode steps: 2370, steps per second: 162, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
301956/2500000: episode: 212, duration: 16.967s, episode steps: 2769, steps per second: 163, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean acti
305232/2500000: episode: 213, duration: 20.135s, episode steps: 3276, steps per second: 163, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
307654/2500000: episode: 214, duration: 15.042s, episode steps: 2422, steps per second: 161, episode reward: -11.000, mean reward: -0.005 [-1.000, 1.000], mean acti
310534/2500000: episode: 215, duration: 17.842s, episode steps: 2880, steps per second: 161, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
313197/2500000: episode: 216, duration: 16.336s, episode steps: 2663, steps per second: 163, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
316107/2500000: episode: 217, duration: 17.782s, episode steps: 2910, steps per second: 164, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean acti
```

```
318624/2500000: episode: 218, duration: 15.542s, episode steps: 2517, steps per second: 162, episode reward: -18.000, mean reward: -0.007 [-1.000, 1.000], mean acti
320766/2500000: episode: 219, duration: 13.397s, episode steps: 2142, steps per second: 160, episode reward: -20.000, mean reward: -0.009 [-1.000, 1.000], mean acti
323156/2500000: episode: 220, duration: 15.097s, episode steps: 2390, steps per second: 158, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
325888/2500000: episode: 221, duration: 17.241s, episode steps: 2732, steps per second: 158, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
327601/2500000: episode: 222, duration: 11.068s, episode steps: 1713, steps per second: 155, episode reward: -19.000, mean reward: -0.011 [-1.000, 1.000], mean acti
330003/2500000: episode: 223, duration: 15.381s, episode steps: 2402, steps per second: 157, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
332167/2500000: episode: 224, duration: 13.569s, episode steps: 2164, steps per second: 159, episode reward: -20.000, mean reward: -0.009 [-1.000, 1.000], mean acti
334940/2500000: episode: 225, duration: 17.804s, episode steps: 2773, steps per second: 156, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean actic
336951/2500000: episode: 226, duration: 12.566s, episode steps: 2011, steps per second: 160, episode reward: -19.000, mean reward: -0.009 [-1.000, 1.000], mean acti
339242/2500000: episode: 227, duration: 14.049s, episode steps: 2291, steps per second: 163, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
341437/2500000: episode: 228, duration: 13.472s, episode steps: 2195, steps per second: 163, episode reward: -19.000, mean reward: -0.009 [-1.000, 1.000], mean acti
343455/2500000: episode: 229, duration: 12.525s, episode steps: 2018, steps per second: 161, episode reward: -19.000, mean reward: -0.009 [-1.000, 1.000], mean acti
347255/2500000: episode: 230, duration: 23.395s, episode steps: 3800, steps per second: 162, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean actic
350611/2500000: episode: 231, duration: 20.631s, episode steps: 3356, steps per second: 163, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
353123/2500000: episode: 232, duration: 15.381s, episode steps: 2512, steps per second: 163, episode reward: -18.000, mean reward: -0.007 [-1.000, 1.000], mean acti
355755/2500000: episode: 233, duration: 16.446s, episode steps: 2632, steps per second: 160, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
359255/2500000: episode: 234, duration: 21.680s, episode steps: 3500, steps per second: 161, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
361836/2500000: episode: 235, duration: 16.065s, episode steps: 2581, steps per second: 161, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
364181/2500000: episode: 236, duration: 14.354s, episode steps: 2345, steps per second: 163, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
366636/2500000: episode: 237, duration: 15.417s, episode steps: 2455, steps per second: 159, episode reward: -19.000, mean reward: -0.008 [-1.000, 1.000], mean acti
369266/2500000: episode: 238, duration: 16.328s, episode steps: 2630, steps per second: 161, episode reward: -17.000, mean reward: -0.006 [-1.000, 1.000], mean acti
371824/2500000: episode: 239, duration: 15.823s, episode steps: 2558, steps per second: 162, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
374948/2500000: episode: 240, duration: 19.327s, episode steps: 3124, steps per second: 162, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
377614/2500000: episode: 241, duration: 17.039s, episode steps: 2666, steps per second: 156, episode reward: -17.000, mean reward: -0.006 [-1.000, 1.000], mean acti
380154/2500000: episode: 242, duration: 15.612s, episode steps: 2540, steps per second: 163, episode reward: -19.000, mean reward: -0.007 [-1.000, 1.000], mean acti
382768/2500000: episode: 243, duration: 16.360s, episode steps: 2614, steps per second: 160, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
386077/2500000: episode: 244, duration: 20.909s, episode steps: 3309, steps per second: 158, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean acti
388705/2500000: episode: 245, duration: 16.006s, episode steps: 2628, steps per second: 164, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
391734/2500000: episode: 246, duration: 18.550s, episode steps: 3029, steps per second: 163, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean acti
394049/2500000: episode: 247, duration: 14.361s, episode steps: 2315, steps per second: 161, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
396319/2500000: episode: 248, duration: 14.354s, episode steps: 2270, steps per second: 158, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
399258/2500000: episode: 249, duration: 18.259s, episode steps: 2939, steps per second: 161, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
401409/2500000: episode: 250, duration: 13.576s, episode steps: 2151, steps per second: 158, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
404361/2500000: episode: 251, duration: 18.528s, episode steps: 2952, steps per second: 159, episode reward: -16.000, mean reward: -0.005 [-1.000, 1.000], mean acti
407076/2500000: episode: 252, duration: 17.079s, episode steps: 2715, steps per second: 159, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
410284/2500000: episode: 253, duration: 20.128s, episode steps: 3208, steps per second: 159, episode reward: -14.000, mean reward: -0.004 [-1.000, 1.000], mean acti
412430/2500000: episode: 254, duration: 13.507s, episode steps: 2146, steps per second: 159, episode reward: -19.000, mean reward: -0.009 [-1.000, 1.000], mean acti
414846/2500000: episode: 255, duration: 14.992s, episode steps: 2416, steps per second: 161, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
417816/2500000: episode: 256, duration: 18.442s, episode steps: 2970, steps per second: 161, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
419466/2500000: episode: 257, duration: 10.408s, episode steps: 1650, steps per second: 159, episode reward: -19.000, mean reward: -0.012 [-1.000, 1.000], mean acti
421760/2500000: episode: 258, duration: 14.438s, episode steps: 2294, steps per second: 159, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
424952/2500000: episode: 259, duration: 20.342s, episode steps: 3192, steps per second: 157, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean acti
428003/2500000: episode: 260, duration: 19.297s, episode steps: 3051, steps per second: 158, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
431779/2500000: episode: 261, duration: 23.743s, episode steps: 3776, steps per second: 159, episode reward: -14.000, mean reward: -0.004 [-1.000, 1.000], mean acti
434988/2500000: episode: 262, duration: 20.346s, episode steps: 3209, steps per second: 158, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean acti
437942/2500000: episode: 263, duration: 18.237s, episode steps: 2954, steps per second: 162, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean actic
440776/2500000: episode: 264, duration: 17.815s, episode steps: 2834, steps per second: 159, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
443966/2500000: episode: 265, duration: 20.005s, episode steps: 3190, steps per second: 159, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
448140/2500000: episode: 266, duration: 26.164s, episode steps: 4174, steps per second: 160, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
450830/2500000: episode: 267, duration: 17.118s, episode steps: 2690, steps per second: 157, episode reward: -16.000, mean reward: -0.006 [-1.000, 1.000], mean acti
453708/2500000: episode: 268, duration: 18.446s, episode steps: 2878, steps per second: 156, episode reward: -16.000, mean reward: -0.006 [-1.000, 1.000], mean acti
456408/2500000: episode: 269, duration: 17.240s, episode steps: 2700, steps per second: 157, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
460019/2500000: episode: 270, duration: 23.014s, episode steps: 3611, steps per second: 157, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
463040/2500000: episode: 271, duration: 18.938s, episode steps: 3021, steps per second: 160, episode reward: -16.000, mean reward: -0.005 [-1.000, 1.000], mean acti
466547/2500000: episode: 272, duration: 21.736s, episode steps: 3507, steps per second: 161, episode reward: -15.000, mean reward: -0.004 [-1.000, 1.000], mean acti
469809/2500000: episode: 273, duration: 20.177s, episode steps: 3262, steps per second: 162, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
472738/2500000: episode: 274, duration: 19.089s, episode steps: 2929, steps per second: 153, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
474663/2500000: episode: 275, duration: 12.319s, episode steps: 1925, steps per second: 156, episode reward: -20.000, mean reward: -0.010 [-1.000, 1.000], mean acti
477284/2500000: episode: 276, duration: 17.030s, episode steps: 2621, steps per second: 154, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
480303/2500000: episode: 277, duration: 19.494s, episode steps: 3019, steps per second: 155, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean actic
482621/2500000: episode: 278, duration: 15.444s, episode steps: 2318, steps per second: 150, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
485764/2500000: episode: 279, duration: 20.177s, episode steps: 3143, steps per second: 156, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean actic
489081/2500000: episode: 280, duration: 20.917s, episode steps: 3317, steps per second: 159, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean actic
491531/2500000: episode: 281, duration: 15.743s, episode steps: 2450, steps per second: 156, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
495161/2500000: episode: 282, duration: 23.273s, episode steps: 3630, steps per second: 156, episode reward: -9.000, mean reward: -0.002 [-1.000, 1.000], mean actic
498970/2500000: episode: 283, duration: 24.292s, episode steps: 3809, steps per second: 157, episode reward: -8.000, mean reward: -0.002 [-1.000, 1.000], mean actic
501175/2500000: episode: 284, duration: 14.093s, episode steps: 2205, steps per second: 156, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
504221/2500000: episode: 285, duration: 19.053s, episode steps: 3046, steps per second: 160, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
507150/2500000: episode: 286, duration: 18.615s, episode steps: 2929, steps per second: 157, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
509423/2500000: episode: 287, duration: 14.693s, episode steps: 2273, steps per second: 155, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
512515/2500000: episode: 288, duration: 19.782s, episode steps: 3092, steps per second: 156, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
515164/2500000: episode: 289, duration: 16.612s, episode steps: 2649, steps per second: 159, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
517689/2500000: episode: 290, duration: 15.701s, episode steps: 2525, steps per second: 161, episode reward: -16.000, mean reward: -0.006 [-1.000, 1.000], mean acti
520181/2500000: episode: 291, duration: 15.747s, episode steps: 2492, steps per second: 158, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
523372/2500000: episode: 292, duration: 20.033s, episode steps: 3191, steps per second: 159, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
527229/2500000: episode: 293, duration: 23.795s, episode steps: 3857, steps per second: 162, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean actic
529946/2500000: episode: 294, duration: 17.429s, episode steps: 2717, steps per second: 156, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
533163/2500000: episode: 295, duration: 20.366s, episode steps: 3217, steps per second: 158, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
537084/2500000: episode: 296, duration: 24.404s, episode steps: 3921, steps per second: 161, episode reward: -8.000, mean reward: -0.002 [-1.000, 1.000], mean actic
540129/2500000: episode: 297, duration: 19.055s, episode steps: 3045, steps per second: 160, episode reward: -17.000, mean reward: -0.006 [-1.000, 1.000], mean acti
543526/2500000: episode: 298, duration: 21.255s, episode steps: 3397, steps per second: 160, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean actic
546263/2500000: episode: 299, duration: 17.106s, episode steps: 2737, steps per second: 160, episode reward: -16.000, mean reward: -0.006 [-1.000, 1.000], mean acti
548708/2500000: episode: 300, duration: 15.222s, episode steps: 2445, steps per second: 161, episode reward: -19.000, mean reward: -0.008 [-1.000, 1.000], mean acti
551747/2500000: episode: 301, duration: 19.535s, episode steps: 3039, steps per second: 156, episode reward: -17.000, mean reward: -0.006 [-1.000, 1.000], mean acti
554682/2500000: episode: 302, duration: 18.704s, episode steps: 2935, steps per second: 157, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
558058/2500000: episode: 303, duration: 21.195s, episode steps: 3376, steps per second: 159, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
560992/2500000: episode: 304, duration: 18.275s, episode steps: 2934, steps per second: 161, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
563917/2500000: episode: 305, duration: 18.161s, episode steps: 2925, steps per second: 161, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
567771/2500000: episode: 306, duration: 24.037s, episode steps: 3854, steps per second: 160, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
570516/2500000: episode: 307, duration: 17.966s, episode steps: 2745, steps per second: 153, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
572994/2500000: episode: 308, duration: 15.639s, episode steps: 2478, steps per second: 158, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
576161/2500000: episode: 309, duration: 19.973s, episode steps: 3167, steps per second: 159, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean actic
578813/2500000: episode: 310, duration: 16.696s, episode steps: 2652, steps per second: 159, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
581906/2500000: episode: 311, duration: 19.914s, episode steps: 3093, steps per second: 155, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
585181/2500000: episode: 312, duration: 20.463s, episode steps: 3275, steps per second: 160, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
588262/2500000: episode: 313, duration: 18.875s, episode steps: 3081, steps per second: 163, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
591490/2500000: episode: 314, duration: 20.400s, episode steps: 3228, steps per second: 158, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
595139/2500000: episode: 315, duration: 23.220s, episode steps: 3649, steps per second: 157, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean actic
597826/2500000: episode: 316, duration: 16.976s, episode steps: 2687, steps per second: 158, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
600592/2500000: episode: 317, duration: 17.464s, episode steps: 2766, steps per second: 158, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
604959/2500000: episode: 318, duration: 27.685s, episode steps: 4367, steps per second: 158, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean actic
609397/2500000: episode: 319, duration: 27.774s, episode steps: 4438, steps per second: 160, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean actic
612496/2500000: episode: 320, duration: 19.163s, episode steps: 3099, steps per second: 162, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
616151/2500000: episode: 321, duration: 23.117s, episode steps: 3655, steps per second: 158, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
619233/2500000: episode: 322, duration: 19.676s, episode steps: 3082, steps per second: 157, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
622635/2500000: episode: 323, duration: 21.325s, episode steps: 3402, steps per second: 160, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean actic
625227/2500000: episode: 324, duration: 16.221s, episode steps: 2592, steps per second: 160, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
627524/2500000: episode: 325, duration: 14.680s, episode steps: 2297, steps per second: 156, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
629350/2500000: episode: 326, duration: 12.143s, episode steps: 1826, steps per second: 150, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
```

```
631370/2500000: episode: 327, duration: 12.778s, episode steps: 2020, steps per second: 158, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
634293/2500000: episode: 328, duration: 18.526s, episode steps: 2923, steps per second: 158, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
636788/2500000: episode: 329, duration: 16.052s, episode steps: 2495, steps per second: 155, episode reward: -16.000, mean reward: -0.006 [-1.000, 1.000], mean acti
639999/2500000: episode: 330, duration: 20.570s, episode steps: 3211, steps per second: 156, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean acti
643974/2500000: episode: 331, duration: 25.573s, episode steps: 3975, steps per second: 155, episode reward: -13.000, mean reward: -0.003 [-1.000, 1.000], mean acti
648300/2500000: episode: 332, duration: 27.685s, episode steps: 4326, steps per second: 156, episode reward: -2.000, mean reward: -0.000 [-1.000, 1.000], mean actio
651763/2500000: episode: 333, duration: 22.459s, episode steps: 3463, steps per second: 154, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
654744/2500000: episode: 334, duration: 19.188s, episode steps: 2981, steps per second: 155, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
657580/2500000: episode: 335, duration: 17.884s, episode steps: 2836, steps per second: 159, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean acti
661565/2500000: episode: 336, duration: 25.064s, episode steps: 3985, steps per second: 159, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
665413/2500000: episode: 337, duration: 24.681s, episode steps: 3848, steps per second: 156, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean acti
669413/2500000: episode: 338, duration: 25.253s, episode steps: 4000, steps per second: 158, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean acti
672893/2500000: episode: 339, duration: 22.079s, episode steps: 3480, steps per second: 158, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
676221/2500000: episode: 340, duration: 21.521s, episode steps: 3328, steps per second: 155, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean acti
680516/2500000: episode: 341, duration: 27.685s, episode steps: 4295, steps per second: 155, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean acti
683295/2500000: episode: 342, duration: 17.663s, episode steps: 2779, steps per second: 157, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
687368/2500000: episode: 343, duration: 25.963s, episode steps: 4073, steps per second: 157, episode reward: -2.000, mean reward: -0.000 [-1.000, 1.000], mean acti
691461/2500000: episode: 344, duration: 26.133s, episode steps: 4093, steps per second: 157, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
694323/2500000: episode: 345, duration: 18.151s, episode steps: 2862, steps per second: 158, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
697951/2500000: episode: 346, duration: 23.139s, episode steps: 3628, steps per second: 157, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean acti
701638/2500000: episode: 347, duration: 23.612s, episode steps: 3687, steps per second: 156, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
704807/2500000: episode: 348, duration: 20.163s, episode steps: 3169, steps per second: 157, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean acti
708449/2500000: episode: 349, duration: 23.356s, episode steps: 3642, steps per second: 156, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean acti
712619/2500000: episode: 350, duration: 26.950s, episode steps: 4170, steps per second: 155, episode reward: -2.000, mean reward: -0.000 [-1.000, 1.000], mean acti
716588/2500000: episode: 351, duration: 25.167s, episode steps: 3969, steps per second: 158, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
719946/2500000: episode: 352, duration: 21.378s, episode steps: 3358, steps per second: 157, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
722573/2500000: episode: 353, duration: 16.673s, episode steps: 2627, steps per second: 158, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
726220/2500000: episode: 354, duration: 23.826s, episode steps: 3647, steps per second: 153, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean acti
729690/2500000: episode: 355, duration: 22.100s, episode steps: 3470, steps per second: 157, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
732783/2500000: episode: 356, duration: 19.882s, episode steps: 3093, steps per second: 156, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
736302/2500000: episode: 357, duration: 22.794s, episode steps: 3519, steps per second: 154, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
740197/2500000: episode: 358, duration: 25.277s, episode steps: 3895, steps per second: 154, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean acti
742639/2500000: episode: 359, duration: 15.840s, episode steps: 2442, steps per second: 154, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
745332/2500000: episode: 360, duration: 17.516s, episode steps: 2693, steps per second: 154, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
748900/2500000: episode: 361, duration: 23.261s, episode steps: 3568, steps per second: 153, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
752717/2500000: episode: 362, duration: 24.650s, episode steps: 3817, steps per second: 155, episode reward: -2.000, mean reward: -0.001 [-1.000, 1.000], mean acti
755380/2500000: episode: 363, duration: 17.054s, episode steps: 2663, steps per second: 156, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
758134/2500000: episode: 364, duration: 17.650s, episode steps: 2754, steps per second: 156, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
761546/2500000: episode: 365, duration: 22.071s, episode steps: 3412, steps per second: 155, episode reward: -8.000, mean reward: -0.002 [-1.000, 1.000], mean acti
764614/2500000: episode: 366, duration: 19.877s, episode steps: 3068, steps per second: 154, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
767720/2500000: episode: 367, duration: 19.752s, episode steps: 3106, steps per second: 157, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
770361/2500000: episode: 368, duration: 16.675s, episode steps: 2641, steps per second: 158, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
772776/2500000: episode: 369, duration: 15.335s, episode steps: 2415, steps per second: 157, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
776178/2500000: episode: 370, duration: 21.910s, episode steps: 3402, steps per second: 155, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
779499/2500000: episode: 371, duration: 20.757s, episode steps: 3321, steps per second: 160, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
782619/2500000: episode: 372, duration: 19.716s, episode steps: 3120, steps per second: 158, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
786091/2500000: episode: 373, duration: 21.860s, episode steps: 3472, steps per second: 159, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
789549/2500000: episode: 374, duration: 21.945s, episode steps: 3458, steps per second: 158, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
793483/2500000: episode: 375, duration: 24.649s, episode steps: 3934, steps per second: 160, episode reward: -2.000, mean reward: -0.001 [-1.000, 1.000], mean acti
796406/2500000: episode: 376, duration: 18.376s, episode steps: 2923, steps per second: 159, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
799892/2500000: episode: 377, duration: 22.079s, episode steps: 3486, steps per second: 158, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean acti
803054/2500000: episode: 378, duration: 20.217s, episode steps: 3162, steps per second: 156, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
806773/2500000: episode: 379, duration: 23.682s, episode steps: 3719, steps per second: 157, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
809762/2500000: episode: 380, duration: 19.227s, episode steps: 2989, steps per second: 155, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean acti
813089/2500000: episode: 381, duration: 21.146s, episode steps: 3327, steps per second: 157, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean acti
816601/2500000: episode: 382, duration: 21.980s, episode steps: 3512, steps per second: 160, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean acti
820405/2500000: episode: 383, duration: 23.537s, episode steps: 3804, steps per second: 162, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean acti
823191/2500000: episode: 384, duration: 17.882s, episode steps: 2786, steps per second: 156, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
826894/2500000: episode: 385, duration: 23.176s, episode steps: 3703, steps per second: 160, episode reward: -8.000, mean reward: -0.002 [-1.000, 1.000], mean acti
829109/2500000: episode: 386, duration: 13.799s, episode steps: 2215, steps per second: 161, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
832471/2500000: episode: 387, duration: 21.332s, episode steps: 3362, steps per second: 158, episode reward: -8.000, mean reward: -0.002 [-1.000, 1.000], mean acti
836045/2500000: episode: 388, duration: 22.272s, episode steps: 3574, steps per second: 160, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
839447/2500000: episode: 389, duration: 21.336s, episode steps: 3402, steps per second: 159, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean acti
842479/2500000: episode: 390, duration: 19.284s, episode steps: 3032, steps per second: 157, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean acti
846088/2500000: episode: 391, duration: 22.878s, episode steps: 3609, steps per second: 158, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean acti
849594/2500000: episode: 392, duration: 22.043s, episode steps: 3506, steps per second: 159, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
852220/2500000: episode: 393, duration: 16.477s, episode steps: 2626, steps per second: 159, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
856057/2500000: episode: 394, duration: 23.790s, episode steps: 3837, steps per second: 161, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean acti
860015/2500000: episode: 395, duration: 24.256s, episode steps: 3958, steps per second: 163, episode reward: -2.000, mean reward: -0.001 [-1.000, 1.000], mean acti
864152/2500000: episode: 396, duration: 25.683s, episode steps: 4137, steps per second: 161, episode reward: 2.000, mean reward: 0.000 [-1.000, 1.000], mean action:
867201/2500000: episode: 397, duration: 18.857s, episode steps: 3049, steps per second: 162, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
869408/2500000: episode: 398, duration: 13.690s, episode steps: 2207, steps per second: 161, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
872947/2500000: episode: 399, duration: 22.658s, episode steps: 3539, steps per second: 156, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
875867/2500000: episode: 400, duration: 18.192s, episode steps: 2920, steps per second: 161, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
879326/2500000: episode: 401, duration: 21.542s, episode steps: 3459, steps per second: 161, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
882575/2500000: episode: 402, duration: 20.370s, episode steps: 3249, steps per second: 159, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
886526/2500000: episode: 403, duration: 25.098s, episode steps: 3951, steps per second: 157, episode reward: -2.000, mean reward: -0.001 [-1.000, 1.000], mean acti
890723/2500000: episode: 404, duration: 26.471s, episode steps: 4197, steps per second: 159, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
893680/2500000: episode: 405, duration: 18.218s, episode steps: 2957, steps per second: 162, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
897473/2500000: episode: 406, duration: 23.748s, episode steps: 3793, steps per second: 160, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean acti
900420/2500000: episode: 407, duration: 18.390s, episode steps: 2947, steps per second: 160, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean acti
903792/2500000: episode: 408, duration: 21.086s, episode steps: 3372, steps per second: 160, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action:
907353/2500000: episode: 409, duration: 22.446s, episode steps: 3561, steps per second: 159, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean acti
910549/2500000: episode: 410, duration: 19.585s, episode steps: 3196, steps per second: 163, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean acti
913708/2500000: episode: 411, duration: 19.671s, episode steps: 3159, steps per second: 161, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
917561/2500000: episode: 412, duration: 23.876s, episode steps: 3853, steps per second: 161, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
919635/2500000: episode: 413, duration: 12.690s, episode steps: 2074, steps per second: 163, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
921891/2500000: episode: 414, duration: 14.363s, episode steps: 2256, steps per second: 157, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
924973/2500000: episode: 415, duration: 18.851s, episode steps: 3082, steps per second: 163, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
928326/2500000: episode: 416, duration: 20.778s, episode steps: 3353, steps per second: 161, episode reward: -8.000, mean reward: -0.002 [-1.000, 1.000], mean acti
931536/2500000: episode: 417, duration: 20.119s, episode steps: 3210, steps per second: 160, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
934920/2500000: episode: 418, duration: 21.018s, episode steps: 3384, steps per second: 161, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
938569/2500000: episode: 419, duration: 22.897s, episode steps: 3649, steps per second: 159, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean acti
941909/2500000: episode: 420, duration: 20.465s, episode steps: 3340, steps per second: 163, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action:
944753/2500000: episode: 421, duration: 17.552s, episode steps: 2844, steps per second: 162, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean actio
948893/2500000: episode: 422, duration: 25.625s, episode steps: 4140, steps per second: 162, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
952597/2500000: episode: 423, duration: 23.300s, episode steps: 3704, steps per second: 159, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean acti
955656/2500000: episode: 424, duration: 18.945s, episode steps: 3059, steps per second: 161, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean acti
959228/2500000: episode: 425, duration: 22.352s, episode steps: 3572, steps per second: 160, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
961901/2500000: episode: 426, duration: 16.704s, episode steps: 2673, steps per second: 160, episode reward: -7.000, mean reward: -0.003 [-1.000, 1.000], mean acti
965431/2500000: episode: 427, duration: 22.373s, episode steps: 3530, steps per second: 158, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
969445/2500000: episode: 428, duration: 25.444s, episode steps: 4014, steps per second: 158, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
973640/2500000: episode: 429, duration: 26.793s, episode steps: 4195, steps per second: 157, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
976615/2500000: episode: 430, duration: 18.513s, episode steps: 2975, steps per second: 161, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
980095/2500000: episode: 431, duration: 21.689s, episode steps: 3480, steps per second: 160, episode reward: -2.000, mean reward: -0.001 [-1.000, 1.000], mean acti
983394/2500000: episode: 432, duration: 20.669s, episode steps: 3299, steps per second: 160, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
986462/2500000: episode: 433, duration: 19.303s, episode steps: 3068, steps per second: 159, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
990037/2500000: episode: 434, duration: 22.113s, episode steps: 3575, steps per second: 162, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
993332/2500000: episode: 435, duration: 20.863s, episode steps: 3295, steps per second: 158, episode reward: -2.000, mean reward: -0.001 [-1.000, 1.000], mean acti
```

```
 997483/2500000: episode: 436, duration: 25.525s, episode steps: 4151, steps per second: 163, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
1001373/2500000: episode: 437, duration: 23.907s, episode steps: 3890, steps per second: 163, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1004963/2500000: episode: 438, duration: 22.329s, episode steps: 3590, steps per second: 161, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1008762/2500000: episode: 439, duration: 23.839s, episode steps: 3799, steps per second: 159, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1012431/2500000: episode: 440, duration: 23.262s, episode steps: 3669, steps per second: 158, episode reward: 9.000, mean reward: 0.002 [-1.000, 1.000], mean actic
1015893/2500000: episode: 441, duration: 21.634s, episode steps: 3462, steps per second: 160, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1019837/2500000: episode: 442, duration: 25.014s, episode steps: 3944, steps per second: 158, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1023577/2500000: episode: 443, duration: 23.590s, episode steps: 3740, steps per second: 159, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1026696/2500000: episode: 444, duration: 19.486s, episode steps: 3119, steps per second: 160, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1030431/2500000: episode: 445, duration: 22.995s, episode steps: 3735, steps per second: 162, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1034668/2500000: episode: 446, duration: 26.535s, episode steps: 4237, steps per second: 160, episode reward: 2.000, mean reward: 0.000 [-1.000, 1.000], mean actic
1038813/2500000: episode: 447, duration: 25.860s, episode steps: 4145, steps per second: 160, episode reward: 6.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1042174/2500000: episode: 448, duration: 21.042s, episode steps: 3361, steps per second: 160, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1046296/2500000: episode: 449, duration: 25.494s, episode steps: 4122, steps per second: 162, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean actic
1049207/2500000: episode: 450, duration: 18.222s, episode steps: 2911, steps per second: 160, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1051981/2500000: episode: 451, duration: 17.367s, episode steps: 2774, steps per second: 160, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1055332/2500000: episode: 452, duration: 20.614s, episode steps: 3351, steps per second: 163, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1059468/2500000: episode: 453, duration: 25.988s, episode steps: 4136, steps per second: 159, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1061985/2500000: episode: 454, duration: 15.721s, episode steps: 2517, steps per second: 160, episode reward: -9.000, mean reward: -0.004 [-1.000, 1.000], mean actic
1065356/2500000: episode: 455, duration: 21.339s, episode steps: 3371, steps per second: 158, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1068948/2500000: episode: 456, duration: 22.769s, episode steps: 3592, steps per second: 158, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1072550/2500000: episode: 457, duration: 22.837s, episode steps: 3602, steps per second: 158, episode reward: -2.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1076598/2500000: episode: 458, duration: 24.890s, episode steps: 4048, steps per second: 163, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1080316/2500000: episode: 459, duration: 23.162s, episode steps: 3718, steps per second: 161, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1083142/2500000: episode: 460, duration: 17.822s, episode steps: 2826, steps per second: 159, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1086293/2500000: episode: 461, duration: 20.026s, episode steps: 3151, steps per second: 157, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean actic
1089644/2500000: episode: 462, duration: 21.337s, episode steps: 3351, steps per second: 157, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1093374/2500000: episode: 463, duration: 23.721s, episode steps: 3730, steps per second: 157, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1096667/2500000: episode: 464, duration: 20.466s, episode steps: 3293, steps per second: 161, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean actic
1100534/2500000: episode: 465, duration: 24.738s, episode steps: 3867, steps per second: 156, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1103745/2500000: episode: 466, duration: 20.517s, episode steps: 3211, steps per second: 157, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1107246/2500000: episode: 467, duration: 22.015s, episode steps: 3501, steps per second: 159, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1110770/2500000: episode: 468, duration: 22.328s, episode steps: 3524, steps per second: 158, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1114029/2500000: episode: 469, duration: 20.378s, episode steps: 3259, steps per second: 160, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1117591/2500000: episode: 470, duration: 22.828s, episode steps: 3562, steps per second: 156, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1119617/2500000: episode: 471, duration: 13.207s, episode steps: 2026, steps per second: 153, episode reward: -9.000, mean reward: -0.004 [-1.000, 1.000], mean actic
1123444/2500000: episode: 472, duration: 24.463s, episode steps: 3827, steps per second: 156, episode reward: -9.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1126814/2500000: episode: 473, duration: 21.830s, episode steps: 3370, steps per second: 154, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean actic
1129985/2500000: episode: 474, duration: 20.175s, episode steps: 3171, steps per second: 157, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1133843/2500000: episode: 475, duration: 24.687s, episode steps: 3858, steps per second: 156, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean actic
1137108/2500000: episode: 476, duration: 21.265s, episode steps: 3265, steps per second: 154, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
1139887/2500000: episode: 477, duration: 17.766s, episode steps: 2779, steps per second: 156, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean actior
1144065/2500000: episode: 478, duration: 27.089s, episode steps: 4178, steps per second: 154, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1147679/2500000: episode: 479, duration: 23.356s, episode steps: 3614, steps per second: 155, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1150256/2500000: episode: 480, duration: 16.479s, episode steps: 2577, steps per second: 156, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1153140/2500000: episode: 481, duration: 18.736s, episode steps: 2884, steps per second: 154, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1156715/2500000: episode: 482, duration: 23.040s, episode steps: 3575, steps per second: 155, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1160171/2500000: episode: 483, duration: 22.371s, episode steps: 3456, steps per second: 154, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean actic
1163730/2500000: episode: 484, duration: 22.910s, episode steps: 3559, steps per second: 155, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1167232/2500000: episode: 485, duration: 22.945s, episode steps: 3502, steps per second: 153, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean actic
1171609/2500000: episode: 486, duration: 28.282s, episode steps: 4377, steps per second: 155, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1174941/2500000: episode: 487, duration: 21.674s, episode steps: 3332, steps per second: 154, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1178350/2500000: episode: 488, duration: 22.039s, episode steps: 3409, steps per second: 155, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1181331/2500000: episode: 489, duration: 19.488s, episode steps: 2981, steps per second: 153, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1184625/2500000: episode: 490, duration: 21.025s, episode steps: 3294, steps per second: 157, episode reward: -8.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1188313/2500000: episode: 491, duration: 23.903s, episode steps: 3688, steps per second: 154, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1191867/2500000: episode: 492, duration: 22.656s, episode steps: 3554, steps per second: 157, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1195559/2500000: episode: 493, duration: 24.136s, episode steps: 3692, steps per second: 153, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1199621/2500000: episode: 494, duration: 26.357s, episode steps: 4062, steps per second: 154, episode reward: -2.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1201811/2500000: episode: 495, duration: 14.150s, episode steps: 2190, steps per second: 155, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
1205273/2500000: episode: 496, duration: 22.613s, episode steps: 3462, steps per second: 153, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean actic
1208648/2500000: episode: 497, duration: 22.249s, episode steps: 3375, steps per second: 152, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1211441/2500000: episode: 498, duration: 18.015s, episode steps: 2793, steps per second: 155, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1214780/2500000: episode: 499, duration: 22.072s, episode steps: 3339, steps per second: 151, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1218915/2500000: episode: 500, duration: 26.713s, episode steps: 4135, steps per second: 155, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1221993/2500000: episode: 501, duration: 19.841s, episode steps: 3078, steps per second: 155, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1225095/2500000: episode: 502, duration: 20.031s, episode steps: 3102, steps per second: 155, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1228704/2500000: episode: 503, duration: 23.462s, episode steps: 3609, steps per second: 154, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1233592/2500000: episode: 504, duration: 31.560s, episode steps: 4888, steps per second: 155, episode reward: 2.000, mean reward: 0.000 [-1.000, 1.000], mean actic
1237355/2500000: episode: 505, duration: 24.113s, episode steps: 3763, steps per second: 156, episode reward: -2.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1241011/2500000: episode: 506, duration: 23.955s, episode steps: 3656, steps per second: 153, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean actic
1244941/2500000: episode: 507, duration: 25.429s, episode steps: 3930, steps per second: 155, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1248736/2500000: episode: 508, duration: 24.737s, episode steps: 3795, steps per second: 153, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1253483/2500000: episode: 509, duration: 30.659s, episode steps: 4747, steps per second: 155, episode reward: -6.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1257125/2500000: episode: 510, duration: 23.723s, episode steps: 3642, steps per second: 154, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
1261046/2500000: episode: 511, duration: 25.749s, episode steps: 3921, steps per second: 152, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1264207/2500000: episode: 512, duration: 20.908s, episode steps: 3161, steps per second: 151, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1267764/2500000: episode: 513, duration: 23.359s, episode steps: 3557, steps per second: 152, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean actic
1270795/2500000: episode: 514, duration: 19.461s, episode steps: 3031, steps per second: 156, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actior
1274699/2500000: episode: 515, duration: 25.728s, episode steps: 3904, steps per second: 152, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1277898/2500000: episode: 516, duration: 20.745s, episode steps: 3199, steps per second: 154, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean actic
1281647/2500000: episode: 517, duration: 24.375s, episode steps: 3749, steps per second: 154, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1284909/2500000: episode: 518, duration: 20.968s, episode steps: 3262, steps per second: 156, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1287584/2500000: episode: 519, duration: 17.657s, episode steps: 2675, steps per second: 151, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1291497/2500000: episode: 520, duration: 25.192s, episode steps: 3913, steps per second: 155, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1295508/2500000: episode: 521, duration: 25.737s, episode steps: 4011, steps per second: 156, episode reward: 2.000, mean reward: 0.000 [-1.000, 1.000], mean actic
1299076/2500000: episode: 522, duration: 22.961s, episode steps: 3568, steps per second: 155, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1302629/2500000: episode: 523, duration: 23.167s, episode steps: 3553, steps per second: 153, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1306393/2500000: episode: 524, duration: 24.382s, episode steps: 3764, steps per second: 154, episode reward: -2.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1310162/2500000: episode: 525, duration: 24.813s, episode steps: 3769, steps per second: 152, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean actic
1314172/2500000: episode: 526, duration: 26.038s, episode steps: 4010, steps per second: 154, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1317257/2500000: episode: 527, duration: 20.077s, episode steps: 3085, steps per second: 154, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1320897/2500000: episode: 528, duration: 23.502s, episode steps: 3640, steps per second: 155, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1323786/2500000: episode: 529, duration: 18.703s, episode steps: 2889, steps per second: 154, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1326903/2500000: episode: 530, duration: 20.348s, episode steps: 3117, steps per second: 153, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1330507/2500000: episode: 531, duration: 23.478s, episode steps: 3604, steps per second: 154, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1333799/2500000: episode: 532, duration: 21.472s, episode steps: 3292, steps per second: 153, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean actic
1338020/2500000: episode: 533, duration: 27.729s, episode steps: 4221, steps per second: 152, episode reward: 2.000, mean reward: 0.000 [-1.000, 1.000], mean actic
1341667/2500000: episode: 534, duration: 23.488s, episode steps: 3647, steps per second: 155, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1345586/2500000: episode: 535, duration: 25.162s, episode steps: 3919, steps per second: 156, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1348645/2500000: episode: 536, duration: 19.669s, episode steps: 3059, steps per second: 156, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean actic
1352660/2500000: episode: 537, duration: 25.975s, episode steps: 4015, steps per second: 155, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean actic
1357118/2500000: episode: 538, duration: 29.357s, episode steps: 4458, steps per second: 152, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1361318/2500000: episode: 539, duration: 27.629s, episode steps: 4200, steps per second: 152, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1365055/2500000: episode: 540, duration: 24.439s, episode steps: 3737, steps per second: 153, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean actic
1368384/2500000: episode: 541, duration: 21.625s, episode steps: 3329, steps per second: 154, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean actic
1371270/2500000: episode: 542, duration: 18.858s, episode steps: 2886, steps per second: 153, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean actior
1375043/2500000: episode: 543, duration: 24.737s, episode steps: 3773, steps per second: 153, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1379807/2500000: episode: 544, duration: 31.541s, episode steps: 4764, steps per second: 151, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
```

```
1383073/2500000: episode: 545, duration: 21.665s, episode steps: 3266, steps per second: 151, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1386147/2500000: episode: 546, duration: 20.181s, episode steps: 3074, steps per second: 152, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actio
1390169/2500000: episode: 547, duration: 26.527s, episode steps: 4022, steps per second: 152, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1394124/2500000: episode: 548, duration: 26.094s, episode steps: 3955, steps per second: 152, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1397325/2500000: episode: 549, duration: 20.915s, episode steps: 3201, steps per second: 153, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1400124/2500000: episode: 550, duration: 18.392s, episode steps: 2799, steps per second: 152, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1403614/2500000: episode: 551, duration: 23.001s, episode steps: 3490, steps per second: 152, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean actio
1406823/2500000: episode: 552, duration: 21.520s, episode steps: 3209, steps per second: 149, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1409744/2500000: episode: 553, duration: 19.190s, episode steps: 2921, steps per second: 152, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1413935/2500000: episode: 554, duration: 27.598s, episode steps: 4191, steps per second: 152, episode reward: -2.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1417718/2500000: episode: 555, duration: 24.926s, episode steps: 3783, steps per second: 152, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1420938/2500000: episode: 556, duration: 21.129s, episode steps: 3220, steps per second: 152, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1424629/2500000: episode: 557, duration: 24.111s, episode steps: 3691, steps per second: 153, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1427601/2500000: episode: 558, duration: 19.595s, episode steps: 2972, steps per second: 152, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1430835/2500000: episode: 559, duration: 21.129s, episode steps: 3234, steps per second: 153, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean actio
1434194/2500000: episode: 560, duration: 22.015s, episode steps: 3359, steps per second: 153, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1437393/2500000: episode: 561, duration: 21.211s, episode steps: 3199, steps per second: 151, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1441245/2500000: episode: 562, duration: 25.186s, episode steps: 3852, steps per second: 153, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1445114/2500000: episode: 563, duration: 25.436s, episode steps: 3869, steps per second: 152, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1448586/2500000: episode: 564, duration: 22.916s, episode steps: 3472, steps per second: 152, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1451726/2500000: episode: 565, duration: 21.001s, episode steps: 3140, steps per second: 150, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1454956/2500000: episode: 566, duration: 21.224s, episode steps: 3230, steps per second: 152, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1457725/2500000: episode: 567, duration: 18.164s, episode steps: 2769, steps per second: 152, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean actio
1461024/2500000: episode: 568, duration: 21.819s, episode steps: 3299, steps per second: 151, episode reward: -8.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1463764/2500000: episode: 569, duration: 18.229s, episode steps: 2740, steps per second: 150, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1467552/2500000: episode: 570, duration: 25.067s, episode steps: 3788, steps per second: 151, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1471413/2500000: episode: 571, duration: 25.894s, episode steps: 3861, steps per second: 149, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1475221/2500000: episode: 572, duration: 25.640s, episode steps: 3808, steps per second: 149, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1479495/2500000: episode: 573, duration: 28.831s, episode steps: 4274, steps per second: 148, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1483267/2500000: episode: 574, duration: 24.925s, episode steps: 3772, steps per second: 151, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1486542/2500000: episode: 575, duration: 21.678s, episode steps: 3275, steps per second: 151, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1490365/2500000: episode: 576, duration: 25.444s, episode steps: 3823, steps per second: 150, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1493249/2500000: episode: 577, duration: 19.341s, episode steps: 2884, steps per second: 149, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1496281/2500000: episode: 578, duration: 20.174s, episode steps: 3032, steps per second: 150, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1499857/2500000: episode: 579, duration: 23.953s, episode steps: 3576, steps per second: 149, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1503809/2500000: episode: 580, duration: 26.291s, episode steps: 3952, steps per second: 150, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1507343/2500000: episode: 581, duration: 23.346s, episode steps: 3534, steps per second: 151, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1510348/2500000: episode: 582, duration: 19.771s, episode steps: 3005, steps per second: 152, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean actio
1513303/2500000: episode: 583, duration: 19.267s, episode steps: 2955, steps per second: 153, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean actio
1516191/2500000: episode: 584, duration: 19.196s, episode steps: 2888, steps per second: 150, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1519147/2500000: episode: 585, duration: 19.437s, episode steps: 2956, steps per second: 152, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1522862/2500000: episode: 586, duration: 24.381s, episode steps: 3715, steps per second: 152, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1526037/2500000: episode: 587, duration: 20.990s, episode steps: 3175, steps per second: 151, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actio
1529476/2500000: episode: 588, duration: 22.587s, episode steps: 3439, steps per second: 152, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1533320/2500000: episode: 589, duration: 25.468s, episode steps: 3844, steps per second: 151, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1536523/2500000: episode: 590, duration: 21.306s, episode steps: 3203, steps per second: 150, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actio
1539410/2500000: episode: 591, duration: 18.967s, episode steps: 2887, steps per second: 152, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1542740/2500000: episode: 592, duration: 21.842s, episode steps: 3330, steps per second: 152, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1546303/2500000: episode: 593, duration: 23.832s, episode steps: 3563, steps per second: 150, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1550576/2500000: episode: 594, duration: 28.240s, episode steps: 4273, steps per second: 151, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1555403/2500000: episode: 595, duration: 31.786s, episode steps: 4827, steps per second: 152, episode reward: -2.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1559018/2500000: episode: 596, duration: 23.825s, episode steps: 3615, steps per second: 152, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1561778/2500000: episode: 597, duration: 18.008s, episode steps: 2760, steps per second: 153, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1564932/2500000: episode: 598, duration: 20.716s, episode steps: 3154, steps per second: 152, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1567964/2500000: episode: 599, duration: 20.074s, episode steps: 3032, steps per second: 151, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actio
1571622/2500000: episode: 600, duration: 24.215s, episode steps: 3658, steps per second: 151, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1575640/2500000: episode: 601, duration: 26.456s, episode steps: 4018, steps per second: 152, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1578825/2500000: episode: 602, duration: 20.861s, episode steps: 3185, steps per second: 153, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1581947/2500000: episode: 603, duration: 20.804s, episode steps: 3122, steps per second: 150, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1585170/2500000: episode: 604, duration: 21.154s, episode steps: 3223, steps per second: 152, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1587522/2500000: episode: 605, duration: 15.337s, episode steps: 2352, steps per second: 153, episode reward: -9.000, mean reward: -0.004 [-1.000, 1.000], mean actic
1590960/2500000: episode: 606, duration: 23.155s, episode steps: 3438, steps per second: 148, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1594580/2500000: episode: 607, duration: 24.216s, episode steps: 3620, steps per second: 149, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1598565/2500000: episode: 608, duration: 26.007s, episode steps: 3985, steps per second: 153, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1601284/2500000: episode: 609, duration: 17.877s, episode steps: 2719, steps per second: 152, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean actio
1604252/2500000: episode: 610, duration: 19.487s, episode steps: 2968, steps per second: 152, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actio
1607799/2500000: episode: 611, duration: 22.982s, episode steps: 3547, steps per second: 154, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1610985/2500000: episode: 612, duration: 20.917s, episode steps: 3186, steps per second: 152, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actio
1614020/2500000: episode: 613, duration: 20.215s, episode steps: 3035, steps per second: 150, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1617111/2500000: episode: 614, duration: 20.362s, episode steps: 3091, steps per second: 152, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1621360/2500000: episode: 615, duration: 28.071s, episode steps: 4249, steps per second: 151, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1624333/2500000: episode: 616, duration: 19.682s, episode steps: 2973, steps per second: 151, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actio
1627422/2500000: episode: 617, duration: 20.188s, episode steps: 3089, steps per second: 153, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1631347/2500000: episode: 618, duration: 25.764s, episode steps: 3925, steps per second: 152, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1633969/2500000: episode: 619, duration: 17.181s, episode steps: 2622, steps per second: 153, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean actio
1637484/2500000: episode: 620, duration: 23.373s, episode steps: 3515, steps per second: 150, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1641458/2500000: episode: 621, duration: 26.332s, episode steps: 3974, steps per second: 151, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1645317/2500000: episode: 622, duration: 25.266s, episode steps: 3859, steps per second: 153, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1649430/2500000: episode: 623, duration: 27.004s, episode steps: 4113, steps per second: 152, episode reward: 2.000, mean reward: 0.000 [-1.000, 1.000], mean action:
1653377/2500000: episode: 624, duration: 25.855s, episode steps: 3947, steps per second: 153, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1656506/2500000: episode: 625, duration: 20.541s, episode steps: 3129, steps per second: 152, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actio
1660244/2500000: episode: 626, duration: 24.567s, episode steps: 3738, steps per second: 152, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1662906/2500000: episode: 627, duration: 17.400s, episode steps: 2662, steps per second: 153, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean actio
1665557/2500000: episode: 628, duration: 17.229s, episode steps: 2651, steps per second: 154, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean actio
1669097/2500000: episode: 629, duration: 23.022s, episode steps: 3540, steps per second: 154, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1672229/2500000: episode: 630, duration: 19.918s, episode steps: 3132, steps per second: 157, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1675690/2500000: episode: 631, duration: 22.332s, episode steps: 3461, steps per second: 155, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1678864/2500000: episode: 632, duration: 20.278s, episode steps: 3174, steps per second: 157, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1681833/2500000: episode: 633, duration: 19.156s, episode steps: 2969, steps per second: 155, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1684713/2500000: episode: 634, duration: 18.292s, episode steps: 2880, steps per second: 157, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1687476/2500000: episode: 635, duration: 18.387s, episode steps: 2763, steps per second: 150, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean actio
1690888/2500000: episode: 636, duration: 21.792s, episode steps: 3412, steps per second: 157, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1693948/2500000: episode: 637, duration: 19.388s, episode steps: 3060, steps per second: 158, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1697157/2500000: episode: 638, duration: 20.214s, episode steps: 3209, steps per second: 159, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actio
1700909/2500000: episode: 639, duration: 23.304s, episode steps: 3752, steps per second: 161, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
1704357/2500000: episode: 640, duration: 21.941s, episode steps: 3448, steps per second: 157, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1707486/2500000: episode: 641, duration: 19.690s, episode steps: 3129, steps per second: 159, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1711004/2500000: episode: 642, duration: 22.022s, episode steps: 3518, steps per second: 160, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1714197/2500000: episode: 643, duration: 20.202s, episode steps: 3193, steps per second: 158, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1718222/2500000: episode: 644, duration: 25.542s, episode steps: 4025, steps per second: 158, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
1720743/2500000: episode: 645, duration: 16.011s, episode steps: 2521, steps per second: 157, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean actio
1723177/2500000: episode: 646, duration: 15.526s, episode steps: 2434, steps per second: 157, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean actio
1726051/2500000: episode: 647, duration: 18.039s, episode steps: 2874, steps per second: 159, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1729473/2500000: episode: 648, duration: 21.961s, episode steps: 3422, steps per second: 156, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1732905/2500000: episode: 649, duration: 21.776s, episode steps: 3432, steps per second: 158, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
1735308/2500000: episode: 650, duration: 15.559s, episode steps: 2403, steps per second: 154, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean actio
1738265/2500000: episode: 651, duration: 18.396s, episode steps: 2957, steps per second: 161, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1742230/2500000: episode: 652, duration: 24.946s, episode steps: 3965, steps per second: 159, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1745645/2500000: episode: 653, duration: 21.596s, episode steps: 3415, steps per second: 158, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
```

```
1748233/2500000: episode: 654, duration: 16.276s, episode steps: 2588, steps per second: 159, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1751923/2500000: episode: 655, duration: 23.422s, episode steps: 3690, steps per second: 158, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1754857/2500000: episode: 656, duration: 18.846s, episode steps: 2934, steps per second: 156, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1757823/2500000: episode: 657, duration: 18.770s, episode steps: 2966, steps per second: 158, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actior
1760456/2500000: episode: 658, duration: 16.521s, episode steps: 2633, steps per second: 159, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean actior
1763583/2500000: episode: 659, duration: 19.877s, episode steps: 3127, steps per second: 157, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1766896/2500000: episode: 660, duration: 21.060s, episode steps: 3313, steps per second: 157, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1770310/2500000: episode: 661, duration: 22.156s, episode steps: 3414, steps per second: 154, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actior
1773182/2500000: episode: 662, duration: 18.334s, episode steps: 2872, steps per second: 157, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean actior
1775819/2500000: episode: 663, duration: 16.888s, episode steps: 2637, steps per second: 156, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean actior
1778430/2500000: episode: 664, duration: 16.501s, episode steps: 2611, steps per second: 158, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean actior
1781181/2500000: episode: 665, duration: 17.608s, episode steps: 2751, steps per second: 156, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean actior
1785020/2500000: episode: 666, duration: 24.924s, episode steps: 3839, steps per second: 154, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1787664/2500000: episode: 667, duration: 16.854s, episode steps: 2644, steps per second: 157, episode reward: 16.000, mean reward: 0.006 [-1.000, 1.000], mean actior
1790930/2500000: episode: 668, duration: 20.728s, episode steps: 3266, steps per second: 158, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1794225/2500000: episode: 669, duration: 21.145s, episode steps: 3295, steps per second: 156, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1797884/2500000: episode: 670, duration: 23.779s, episode steps: 3659, steps per second: 154, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1801050/2500000: episode: 671, duration: 21.042s, episode steps: 3166, steps per second: 150, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1803947/2500000: episode: 672, duration: 18.957s, episode steps: 2897, steps per second: 153, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean actior
1807164/2500000: episode: 673, duration: 21.058s, episode steps: 3217, steps per second: 153, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1810031/2500000: episode: 674, duration: 18.488s, episode steps: 2867, steps per second: 155, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actior
1812819/2500000: episode: 675, duration: 18.208s, episode steps: 2788, steps per second: 153, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean actior
1816324/2500000: episode: 676, duration: 22.690s, episode steps: 3505, steps per second: 154, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1819462/2500000: episode: 677, duration: 20.081s, episode steps: 3138, steps per second: 156, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1821951/2500000: episode: 678, duration: 16.064s, episode steps: 2489, steps per second: 155, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean actior
1824876/2500000: episode: 679, duration: 18.595s, episode steps: 2925, steps per second: 157, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1827398/2500000: episode: 680, duration: 15.937s, episode steps: 2522, steps per second: 158, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean actior
1830468/2500000: episode: 681, duration: 19.614s, episode steps: 3070, steps per second: 157, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1833539/2500000: episode: 682, duration: 20.098s, episode steps: 3071, steps per second: 153, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1837060/2500000: episode: 683, duration: 22.321s, episode steps: 3521, steps per second: 158, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1839904/2500000: episode: 684, duration: 17.901s, episode steps: 2844, steps per second: 159, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1842489/2500000: episode: 685, duration: 16.559s, episode steps: 2585, steps per second: 156, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean actior
1845090/2500000: episode: 686, duration: 16.427s, episode steps: 2601, steps per second: 158, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1847569/2500000: episode: 687, duration: 15.868s, episode steps: 2479, steps per second: 156, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean actior
1850479/2500000: episode: 688, duration: 18.253s, episode steps: 2910, steps per second: 159, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1853360/2500000: episode: 689, duration: 17.985s, episode steps: 2881, steps per second: 160, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1855658/2500000: episode: 690, duration: 14.321s, episode steps: 2298, steps per second: 160, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean actior
1857953/2500000: episode: 691, duration: 14.426s, episode steps: 2295, steps per second: 159, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean actior
1860051/2500000: episode: 692, duration: 13.108s, episode steps: 2098, steps per second: 160, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean actior
1862439/2500000: episode: 693, duration: 15.019s, episode steps: 2388, steps per second: 159, episode reward: 11.000, mean reward: 0.005 [-1.000, 1.000], mean actior
1865192/2500000: episode: 694, duration: 17.255s, episode steps: 2753, steps per second: 160, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean actior
1867338/2500000: episode: 695, duration: 13.418s, episode steps: 2146, steps per second: 160, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean actior
1869952/2500000: episode: 696, duration: 16.590s, episode steps: 2614, steps per second: 158, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean actior
1872832/2500000: episode: 697, duration: 17.985s, episode steps: 2880, steps per second: 160, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean actior
1876616/2500000: episode: 698, duration: 23.660s, episode steps: 3784, steps per second: 160, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1879178/2500000: episode: 699, duration: 15.951s, episode steps: 2562, steps per second: 161, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean actior
1881653/2500000: episode: 700, duration: 15.971s, episode steps: 2475, steps per second: 155, episode reward: 9.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1884616/2500000: episode: 701, duration: 18.378s, episode steps: 2963, steps per second: 161, episode reward: 13.000, mean reward: 0.004 [-1.000, 1.000], mean actior
1887408/2500000: episode: 702, duration: 17.293s, episode steps: 2792, steps per second: 161, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1890834/2500000: episode: 703, duration: 21.286s, episode steps: 3426, steps per second: 161, episode reward: -2.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1893854/2500000: episode: 704, duration: 18.629s, episode steps: 3020, steps per second: 162, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1896785/2500000: episode: 705, duration: 18.510s, episode steps: 2931, steps per second: 158, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1899095/2500000: episode: 706, duration: 14.385s, episode steps: 2310, steps per second: 161, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean actior
1902010/2500000: episode: 707, duration: 17.942s, episode steps: 2915, steps per second: 162, episode reward: 15.000, mean reward: 0.005 [-1.000, 1.000], mean actior
1905731/2500000: episode: 708, duration: 23.050s, episode steps: 3721, steps per second: 161, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1907680/2500000: episode: 709, duration: 12.182s, episode steps: 1949, steps per second: 160, episode reward: 19.000, mean reward: 0.010 [-1.000, 1.000], mean actior
1910599/2500000: episode: 710, duration: 18.203s, episode steps: 2919, steps per second: 160, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1913509/2500000: episode: 711, duration: 18.146s, episode steps: 2910, steps per second: 160, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1916187/2500000: episode: 712, duration: 16.924s, episode steps: 2678, steps per second: 158, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean actior
1918739/2500000: episode: 713, duration: 15.956s, episode steps: 2552, steps per second: 160, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean actior
1921053/2500000: episode: 714, duration: 14.234s, episode steps: 2314, steps per second: 163, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean actior
1923959/2500000: episode: 715, duration: 17.935s, episode steps: 2906, steps per second: 162, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actior
1926906/2500000: episode: 716, duration: 18.126s, episode steps: 2947, steps per second: 163, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1929438/2500000: episode: 717, duration: 15.880s, episode steps: 2532, steps per second: 159, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean actior
1932180/2500000: episode: 718, duration: 17.393s, episode steps: 2742, steps per second: 158, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean actior
1934836/2500000: episode: 719, duration: 16.176s, episode steps: 2656, steps per second: 164, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean actior
1937375/2500000: episode: 720, duration: 15.782s, episode steps: 2539, steps per second: 161, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean actior
1940175/2500000: episode: 721, duration: 17.414s, episode steps: 2800, steps per second: 161, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean actior
1942622/2500000: episode: 722, duration: 15.255s, episode steps: 2447, steps per second: 160, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean actior
1945043/2500000: episode: 723, duration: 15.210s, episode steps: 2421, steps per second: 159, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean actior
1947453/2500000: episode: 724, duration: 15.236s, episode steps: 2410, steps per second: 158, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean actior
1950613/2500000: episode: 725, duration: 19.703s, episode steps: 3160, steps per second: 160, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1952883/2500000: episode: 726, duration: 14.234s, episode steps: 2270, steps per second: 159, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean actior
1955253/2500000: episode: 727, duration: 14.552s, episode steps: 2370, steps per second: 163, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean actior
1958913/2500000: episode: 728, duration: 22.593s, episode steps: 3660, steps per second: 162, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1961733/2500000: episode: 729, duration: 17.218s, episode steps: 2820, steps per second: 164, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean actior
1964766/2500000: episode: 730, duration: 18.760s, episode steps: 3033, steps per second: 162, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean actior
1967800/2500000: episode: 731, duration: 18.785s, episode steps: 3034, steps per second: 162, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1969986/2500000: episode: 732, duration: 13.293s, episode steps: 2186, steps per second: 164, episode reward: 19.000, mean reward: 0.009 [-1.000, 1.000], mean actior
1971999/2500000: episode: 733, duration: 12.568s, episode steps: 2013, steps per second: 160, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean actior
1974892/2500000: episode: 734, duration: 17.698s, episode steps: 2893, steps per second: 163, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actior
1978026/2500000: episode: 735, duration: 19.492s, episode steps: 3134, steps per second: 161, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean actic
1980220/2500000: episode: 736, duration: 13.743s, episode steps: 2194, steps per second: 160, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean actior
1982821/2500000: episode: 737, duration: 16.145s, episode steps: 2601, steps per second: 161, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean actior
1985165/2500000: episode: 738, duration: 14.507s, episode steps: 2344, steps per second: 162, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean actior
1988435/2500000: episode: 739, duration: 20.385s, episode steps: 3270, steps per second: 160, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1990786/2500000: episode: 740, duration: 14.397s, episode steps: 2351, steps per second: 163, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean actior
1994282/2500000: episode: 741, duration: 21.703s, episode steps: 3496, steps per second: 161, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1998264/2500000: episode: 742, duration: 24.596s, episode steps: 3982, steps per second: 162, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2001092/2500000: episode: 743, duration: 17.398s, episode steps: 2828, steps per second: 163, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean actior
2004663/2500000: episode: 744, duration: 22.102s, episode steps: 3571, steps per second: 162, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
2007654/2500000: episode: 745, duration: 18.492s, episode steps: 2991, steps per second: 162, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean actior
2009487/2500000: episode: 746, duration: 11.244s, episode steps: 1833, steps per second: 163, episode reward: 18.000, mean reward: 0.010 [-1.000, 1.000], mean actior
2012284/2500000: episode: 747, duration: 17.153s, episode steps: 2797, steps per second: 163, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean actior
2015978/2500000: episode: 748, duration: 23.135s, episode steps: 3694, steps per second: 160, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actior
2018160/2500000: episode: 749, duration: 13.574s, episode steps: 2182, steps per second: 161, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean actior
2021845/2500000: episode: 750, duration: 22.479s, episode steps: 3685, steps per second: 164, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
2024204/2500000: episode: 751, duration: 14.634s, episode steps: 2359, steps per second: 161, episode reward: 11.000, mean reward: 0.005 [-1.000, 1.000], mean actior
2026893/2500000: episode: 752, duration: 16.484s, episode steps: 2689, steps per second: 163, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean actior
2030073/2500000: episode: 753, duration: 20.142s, episode steps: 3180, steps per second: 158, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2033268/2500000: episode: 754, duration: 19.682s, episode steps: 3195, steps per second: 162, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2036518/2500000: episode: 755, duration: 20.189s, episode steps: 3250, steps per second: 161, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2039509/2500000: episode: 756, duration: 18.700s, episode steps: 2991, steps per second: 160, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
2041999/2500000: episode: 757, duration: 15.505s, episode steps: 2490, steps per second: 161, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean actior
2044152/2500000: episode: 758, duration: 13.567s, episode steps: 2153, steps per second: 159, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean actior
2046949/2500000: episode: 759, duration: 17.333s, episode steps: 2797, steps per second: 161, episode reward: 7.000, mean reward: 0.003 [-1.000, 1.000], mean action:
2049150/2500000: episode: 760, duration: 13.623s, episode steps: 2201, steps per second: 162, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean actior
2051139/2500000: episode: 761, duration: 12.326s, episode steps: 1989, steps per second: 161, episode reward: 17.000, mean reward: 0.009 [-1.000, 1.000], mean actior
2053181/2500000: episode: 762, duration: 12.746s, episode steps: 2042, steps per second: 160, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean actior
```

```
2055617/2500000: episode: 763, duration: 14.883s, episode steps: 2436, steps per second: 164, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean action
2058813/2500000: episode: 764, duration: 19.757s, episode steps: 3196, steps per second: 162, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2061078/2500000: episode: 765, duration: 14.365s, episode steps: 2265, steps per second: 158, episode reward: 16.000, mean reward: 0.007 [-1.000, 1.000], mean action
2063491/2500000: episode: 766, duration: 15.085s, episode steps: 2413, steps per second: 160, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean action
2066571/2500000: episode: 767, duration: 19.168s, episode steps: 3080, steps per second: 161, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2069584/2500000: episode: 768, duration: 18.606s, episode steps: 3013, steps per second: 162, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean action
2073669/2500000: episode: 769, duration: 25.295s, episode steps: 4085, steps per second: 161, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
2075931/2500000: episode: 770, duration: 14.240s, episode steps: 2262, steps per second: 159, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean action
2078718/2500000: episode: 771, duration: 17.483s, episode steps: 2787, steps per second: 159, episode reward: 7.000, mean reward: 0.003 [-1.000, 1.000], mean action:
2080990/2500000: episode: 772, duration: 14.792s, episode steps: 2272, steps per second: 154, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean action
2082840/2500000: episode: 773, duration: 11.965s, episode steps: 1850, steps per second: 155, episode reward: 18.000, mean reward: 0.010 [-1.000, 1.000], mean action
2085161/2500000: episode: 774, duration: 14.560s, episode steps: 2321, steps per second: 159, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean action
2087419/2500000: episode: 775, duration: 14.076s, episode steps: 2258, steps per second: 160, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action
2090239/2500000: episode: 776, duration: 17.879s, episode steps: 2820, steps per second: 158, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action
2092415/2500000: episode: 777, duration: 13.812s, episode steps: 2176, steps per second: 158, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean action
2094918/2500000: episode: 778, duration: 15.931s, episode steps: 2503, steps per second: 157, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action
2097735/2500000: episode: 779, duration: 17.769s, episode steps: 2817, steps per second: 159, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean action
2100016/2500000: episode: 780, duration: 14.170s, episode steps: 2281, steps per second: 161, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean action
2102530/2500000: episode: 781, duration: 15.808s, episode steps: 2514, steps per second: 159, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action
2104582/2500000: episode: 782, duration: 13.053s, episode steps: 2052, steps per second: 157, episode reward: 16.000, mean reward: 0.008 [-1.000, 1.000], mean action
2106780/2500000: episode: 783, duration: 13.730s, episode steps: 2198, steps per second: 160, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean action
2109652/2500000: episode: 784, duration: 18.487s, episode steps: 2872, steps per second: 155, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
2112507/2500000: episode: 785, duration: 17.994s, episode steps: 2855, steps per second: 159, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean action
2115025/2500000: episode: 786, duration: 15.994s, episode steps: 2518, steps per second: 157, episode reward: 16.000, mean reward: 0.006 [-1.000, 1.000], mean action
2118465/2500000: episode: 787, duration: 21.662s, episode steps: 3440, steps per second: 159, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2120481/2500000: episode: 788, duration: 12.735s, episode steps: 2016, steps per second: 158, episode reward: 17.000, mean reward: 0.008 [-1.000, 1.000], mean action
2123135/2500000: episode: 789, duration: 16.793s, episode steps: 2654, steps per second: 158, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean action
2125645/2500000: episode: 790, duration: 15.678s, episode steps: 2510, steps per second: 160, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean action
2128797/2500000: episode: 791, duration: 20.437s, episode steps: 3152, steps per second: 154, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
2131715/2500000: episode: 792, duration: 18.317s, episode steps: 2918, steps per second: 159, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean action
2134636/2500000: episode: 793, duration: 18.379s, episode steps: 2921, steps per second: 159, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
2136687/2500000: episode: 794, duration: 13.147s, episode steps: 2051, steps per second: 156, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean action
2141057/2500000: episode: 795, duration: 28.115s, episode steps: 4370, steps per second: 155, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean action
2143879/2500000: episode: 796, duration: 18.103s, episode steps: 2822, steps per second: 156, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean action
2146631/2500000: episode: 797, duration: 17.680s, episode steps: 2752, steps per second: 156, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean action
```