```
Training for 1500000 steps ...
    822/1500000: episode: 1, duration: 7.362s, episode steps: 822, steps per second: 112, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action:
   1646/1500000: episode: 2, duration: 1.897s, episode steps: 824, steps per second: 434, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
   2550/1500000: episode: 3, duration: 2.017s, episode steps: 904, steps per second: 448, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action:
   3430/1500000: episode: 4, duration: 1.993s, episode steps: 880, steps per second: 442, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
   4194/1500000: episode: 5, duration: 1.754s, episode steps: 764, steps per second: 436, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
   5064/1500000: episode: 6, duration: 1.938s, episode steps: 870, steps per second: 449, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action:
   5888/1500000: episode: 7, duration: 1.861s, episode steps: 824, steps per second: 443, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
   6868/1500000: episode: 8, duration: 2.200s, episode steps: 980, steps per second: 446, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action:
   7688/1500000: episode: 9, duration: 1.846s, episode steps: 820, steps per second: 444, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action:
   8452/1500000: episode: 10, duration: 1.730s, episode steps: 764, steps per second: 442, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
   9216/1500000: episode: 11, duration: 1.727s, episode steps: 764, steps per second: 442, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
   9980/1500000: episode: 12, duration: 1.687s, episode steps: 764, steps per second: 453, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
  10852/1500000: episode: 13, duration: 5.477s, episode steps: 872, steps per second: 159, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action:
  11870/1500000: episode: 14, duration: 5.610s, episode steps: 1018, steps per second: 181, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action
  12752/1500000: episode: 15, duration: 4.888s, episode steps: 882, steps per second: 180, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
  13789/1500000: episode: 16, duration: 5.739s, episode steps: 1037, steps per second: 181, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action
  14673/1500000: episode: 17, duration: 5.001s, episode steps: 884, steps per second: 177, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
  15578/1500000: episode: 18, duration: 5.043s, episode steps: 905, steps per second: 179, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action:
  16399/1500000: episode: 19, duration: 4.659s, episode steps: 821, steps per second: 176, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action:
  17303/1500000: episode: 20, duration: 5.158s, episode steps: 904, steps per second: 175, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action:
  18174/1500000: episode: 21, duration: 4.945s, episode steps: 871, steps per second: 176, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
  19000/1500000: episode: 22, duration: 4.586s, episode steps: 826, steps per second: 180, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
  20120/1500000: episode: 23, duration: 6.188s, episode steps: 1120, steps per second: 181, episode reward: -18.000, mean reward: -0.016 [-1.000, 1.000], mean action
  21029/1500000: episode: 24, duration: 5.031s, episode steps: 909, steps per second: 181, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action:
  21872/1500000: episode: 25, duration: 4.743s, episode steps: 843, steps per second: 178, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
  22714/1500000: episode: 26, duration: 4.650s, episode steps: 842, steps per second: 181, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action:
  23602/1500000: episode: 27, duration: 5.002s, episode steps: 888, steps per second: 178, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
  24682/1500000: episode: 28, duration: 5.898s, episode steps: 1080, steps per second: 183, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action
  25944/1500000: episode: 29, duration: 7.065s, episode steps: 1262, steps per second: 179, episode reward: -17.000, mean reward: -0.013 [-1.000, 1.000], mean action
  26877/1500000: episode: 30, duration: 5.235s, episode steps: 933, steps per second: 178, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action:
  27821/1500000: episode: 31, duration: 5.314s, episode steps: 944, steps per second: 178, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action:
  28691/1500000: episode: 32, duration: 4.819s, episode steps: 870, steps per second: 181, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action:
  29562/1500000: episode: 33, duration: 4.858s, episode steps: 871, steps per second: 179, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action:
  30587/1500000: episode: 34, duration: 5.624s, episode steps: 1025, steps per second: 182, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action
  31457/1500000: episode: 35, duration: 4.862s, episode steps: 870, steps per second: 179, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action:
  32281/1500000: episode: 36, duration: 4.596s, episode steps: 824, steps per second: 179, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
  33161/1500000: episode: 37, duration: 4.850s, episode steps: 880, steps per second: 181, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
  33925/1500000: episode: 38, duration: 4.143s, episode steps: 764, steps per second: 184, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
  34846/1500000: episode: 39, duration: 5.108s, episode steps: 921, steps per second: 180, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action:
  35887/1500000: episode: 40, duration: 5.683s, episode steps: 1041, steps per second: 183, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action
  36651/1500000: episode: 41, duration: 4.180s, episode steps: 764, steps per second: 183, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
  37723/1500000: episode: 42, duration: 5.747s, episode steps: 1072, steps per second: 187, episode reward: -19.000, mean reward: -0.018 [-1.000, 1.000], mean action
  38547/1500000: episode: 43, duration: 4.513s, episode steps: 824, steps per second: 183, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
  39409/1500000: episode: 44, duration: 4.653s, episode steps: 862, steps per second: 185, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action:
  40234/1500000: episode: 45, duration: 4.474s, episode steps: 825, steps per second: 184, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
  40998/1500000: episode: 46, duration: 4.125s, episode steps: 764, steps per second: 185, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
  41762/1500000: episode: 47, duration: 4.120s, episode steps: 764, steps per second: 185, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
  42582/1500000: episode: 48, duration: 4.437s, episode steps: 820, steps per second: 185, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action:
  43520/1500000: episode: 49, duration: 5.130s, episode steps: 938, steps per second: 183, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action:
  44530/1500000: episode: 50, duration: 5.565s, episode steps: 1010, steps per second: 181, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action
  45322/1500000: episode: 51, duration: 4.331s, episode steps: 792, steps per second: 183, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
  46105/1500000: episode: 52, duration: 4.254s, episode steps: 783, steps per second: 184, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
  47025/1500000: episode: 53, duration: 4.978s, episode steps: 920, steps per second: 185, episode reward: -19.000, mean reward: -0.021 [-1.000, 1.000], mean action:
  47878/1500000: episode: 54, duration: 4.637s, episode steps: 853, steps per second: 184, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
  48766/1500000: episode: 55, duration: 4.962s, episode steps: 888, steps per second: 179, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
  49868/1500000: episode: 56, duration: 6.059s, episode steps: 1102, steps per second: 182, episode reward: -19.000, mean reward: -0.017 [-1.000, 1.000], mean action
  50800/1500000: episode: 57, duration: 5.120s, episode steps: 932, steps per second: 182, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action:
  51652/1500000: episode: 58, duration: 4.676s, episode steps: 852, steps per second: 182, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
  52600/1500000: episode: 59, duration: 5.225s, episode steps: 948, steps per second: 181, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action:
  53424/1500000: episode: 60, duration: 4.535s, episode steps: 824, steps per second: 182, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
  54356/1500000: episode: 61, duration: 5.160s, episode steps: 932, steps per second: 181, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action:
  55268/1500000: episode: 62, duration: 5.051s, episode steps: 912, steps per second: 181, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action:
  56093/1500000: episode: 63, duration: 4.571s, episode steps: 825, steps per second: 180, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
  57005/1500000: episode: 64, duration: 5.073s, episode steps: 912, steps per second: 180, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action:
  57871/1500000: episode: 65, duration: 4.784s, episode steps: 866, steps per second: 181, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action:
  58695/1500000: episode: 66, duration: 4.582s, episode steps: 824, steps per second: 180, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
  59547/1500000: episode: 67, duration: 4.813s, episode steps: 852, steps per second: 177, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
  60466/1500000: episode: 68, duration: 5.096s, episode steps: 919, steps per second: 180, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action:
  61290/1500000: episode: 69, duration: 4.496s, episode steps: 824, steps per second: 183, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
  62135/1500000: episode: 70, duration: 4.649s, episode steps: 845, steps per second: 182, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
  62978/1500000: episode: 71, duration: 4.630s, episode steps: 843, steps per second: 182, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action:
  63938/1500000: episode: 72, duration: 5.302s, episode steps: 960, steps per second: 181, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action:
  64854/1500000: episode: 73, duration: 5.038s, episode steps: 916, steps per second: 182, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action:
  65888/1500000: episode: 74, duration: 5.742s, episode steps: 1034, steps per second: 180, episode reward: -21.000, mean reward: -0.020 [-1.000, 0.000], mean action
  66732/1500000: episode: 75, duration: 4.588s, episode steps: 844, steps per second: 184, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action:
  67781/1500000: episode: 76, duration: 5.790s, episode steps: 1049, steps per second: 181, episode reward: -18.000, mean reward: -0.017 [-1.000, 1.000], mean action
  68592/1500000: episode: 77, duration: 4.427s, episode steps: 811, steps per second: 183, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action:
  69432/1500000: episode: 78, duration: 4.586s, episode steps: 840, steps per second: 183, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
  70196/1500000: episode: 79, duration: 4.362s, episode steps: 764, steps per second: 175, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
  71010/1500000: episode: 80, duration: 4.519s, episode steps: 814, steps per second: 180, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action:
  71834/1500000: episode: 81, duration: 4.566s, episode steps: 824, steps per second: 180, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
  72672/1500000: episode: 82, duration: 4.830s, episode steps: 838, steps per second: 173, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action:
  73514/1500000: episode: 83, duration: 4.874s, episode steps: 842, steps per second: 173, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action:
  74278/1500000: episode: 84, duration: 4.201s, episode steps: 764, steps per second: 182, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
  75139/1500000: episode: 85, duration: 4.644s, episode steps: 861, steps per second: 185, episode reward: -20.000, mean reward: -0.023 [-1.000, 0.000], mean action:
  75903/1500000: episode: 86, duration: 4.182s, episode steps: 764, steps per second: 183, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
  76748/1500000: episode: 87, duration: 4.759s, episode steps: 845, steps per second: 178, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
  77725/1500000: episode: 88, duration: 5.409s, episode steps: 977, steps per second: 181, episode reward: -20.000, mean reward: -0.020 [-1.000, 0.000], mean action:
  78566/1500000: episode: 89, duration: 4.573s, episode steps: 841, steps per second: 184, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
  79358/1500000: episode: 90, duration: 4.310s, episode steps: 792, steps per second: 184, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
  80216/1500000: episode: 91, duration: 4.749s, episode steps: 858, steps per second: 181, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
  80980/1500000: episode: 92, duration: 4.192s, episode steps: 764, steps per second: 182, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
  81822/1500000: episode: 93, duration: 4.676s, episode steps: 842, steps per second: 180, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
  82854/1500000: episode: 94, duration: 5.618s, episode steps: 1032, steps per second: 184, episode reward: -18.000, mean reward: -0.017 [-1.000, 1.000], mean action
  83825/1500000: episode: 95, duration: 5.266s, episode steps: 971, steps per second: 184, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action:
  84729/1500000: episode: 96, duration: 4.893s, episode steps: 904, steps per second: 185, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action:
  85571/1500000: episode: 97, duration: 4.597s, episode steps: 842, steps per second: 183, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action:
  86556/1500000: episode: 98, duration: 5.421s, episode steps: 985, steps per second: 182, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action:
  87394/1500000: episode: 99, duration: 4.684s, episode steps: 838, steps per second: 179, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action:
  88186/1500000: episode: 100, duration: 4.419s, episode steps: 792, steps per second: 179, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action
  89091/1500000: episode: 101, duration: 4.980s, episode steps: 905, steps per second: 182, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action
  89855/1500000: episode: 102, duration: 4.135s, episode steps: 764, steps per second: 185, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action
  90725/1500000: episode: 103, duration: 4.762s, episode steps: 870, steps per second: 183, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action
  91507/1500000: episode: 104, duration: 4.316s, episode steps: 782, steps per second: 181, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action
  92387/1500000: episode: 105, duration: 4.835s, episode steps: 880, steps per second: 182, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action
  93317/1500000: episode: 106, duration: 5.096s, episode steps: 930, steps per second: 182, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action
  94159/1500000: episode: 107, duration: 4.632s, episode steps: 842, steps per second: 182, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action
  95109/1500000: episode: 108, duration: 5.164s, episode steps: 950, steps per second: 184, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action
```

```
 96008/1500000: episode: 109, duration: 4.897s, episode steps: 899, steps per second: 184, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action
 97346/1500000: episode: 110, duration: 7.241s, episode steps: 1338, steps per second: 185, episode reward: -21.000, mean reward: -0.016 [-1.000, 0.000], mean actio
 98294/1500000: episode: 111, duration: 5.154s, episode steps: 948, steps per second: 184, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action
 99173/1500000: episode: 112, duration: 4.788s, episode steps: 879, steps per second: 184, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action
100087/1500000: episode: 113, duration: 5.028s, episode steps: 914, steps per second: 182, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action
100879/1500000: episode: 114, duration: 4.320s, episode steps: 792, steps per second: 183, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action
101828/1500000: episode: 115, duration: 5.258s, episode steps: 949, steps per second: 180, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action
102712/1500000: episode: 116, duration: 4.889s, episode steps: 884, steps per second: 181, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action
103678/1500000: episode: 117, duration: 5.341s, episode steps: 966, steps per second: 181, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action
104461/1500000: episode: 118, duration: 4.276s, episode steps: 783, steps per second: 183, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action
105313/1500000: episode: 119, duration: 4.665s, episode steps: 852, steps per second: 183, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
106239/1500000: episode: 120, duration: 5.055s, episode steps: 926, steps per second: 183, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action
107023/1500000: episode: 121, duration: 4.236s, episode steps: 784, steps per second: 185, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action
107806/1500000: episode: 122, duration: 4.220s, episode steps: 783, steps per second: 186, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action
108649/1500000: episode: 123, duration: 4.616s, episode steps: 843, steps per second: 183, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
109638/1500000: episode: 124, duration: 5.404s, episode steps: 989, steps per second: 183, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action
110430/1500000: episode: 125, duration: 4.344s, episode steps: 792, steps per second: 182, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action
111365/1500000: episode: 126, duration: 5.064s, episode steps: 935, steps per second: 185, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action
112147/1500000: episode: 127, duration: 4.317s, episode steps: 782, steps per second: 181, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action
113110/1500000: episode: 128, duration: 5.244s, episode steps: 963, steps per second: 184, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action
113948/1500000: episode: 129, duration: 4.584s, episode steps: 838, steps per second: 183, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action
114815/1500000: episode: 130, duration: 4.782s, episode steps: 867, steps per second: 181, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action
115701/1500000: episode: 131, duration: 4.814s, episode steps: 886, steps per second: 184, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action
116539/1500000: episode: 132, duration: 4.760s, episode steps: 838, steps per second: 176, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action
117322/1500000: episode: 133, duration: 4.413s, episode steps: 783, steps per second: 177, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action
118266/1500000: episode: 134, duration: 5.119s, episode steps: 944, steps per second: 184, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action
119210/1500000: episode: 135, duration: 5.227s, episode steps: 944, steps per second: 181, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action
120034/1500000: episode: 136, duration: 4.568s, episode steps: 824, steps per second: 180, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
120817/1500000: episode: 137, duration: 4.269s, episode steps: 783, steps per second: 183, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action
121660/1500000: episode: 138, duration: 4.657s, episode steps: 843, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
122544/1500000: episode: 139, duration: 4.940s, episode steps: 884, steps per second: 179, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action
123490/1500000: episode: 140, duration: 5.402s, episode steps: 946, steps per second: 175, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action
124314/1500000: episode: 141, duration: 4.521s, episode steps: 824, steps per second: 182, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
125157/1500000: episode: 142, duration: 4.881s, episode steps: 843, steps per second: 173, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
125981/1500000: episode: 143, duration: 4.540s, episode steps: 824, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
126893/1500000: episode: 144, duration: 5.037s, episode steps: 912, steps per second: 181, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action
127717/1500000: episode: 145, duration: 4.552s, episode steps: 824, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
128541/1500000: episode: 146, duration: 4.652s, episode steps: 824, steps per second: 177, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
129365/1500000: episode: 147, duration: 4.730s, episode steps: 824, steps per second: 174, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
130189/1500000: episode: 148, duration: 4.708s, episode steps: 824, steps per second: 175, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
131013/1500000: episode: 149, duration: 4.649s, episode steps: 824, steps per second: 177, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
131839/1500000: episode: 150, duration: 4.739s, episode steps: 826, steps per second: 174, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
132843/1500000: episode: 151, duration: 5.549s, episode steps: 1004, steps per second: 181, episode reward: -21.000, mean reward: -0.021 [-1.000, 0.000], mean actio
133667/1500000: episode: 152, duration: 4.550s, episode steps: 824, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
134450/1500000: episode: 153, duration: 4.303s, episode steps: 783, steps per second: 182, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action
135274/1500000: episode: 154, duration: 4.538s, episode steps: 824, steps per second: 182, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
136195/1500000: episode: 155, duration: 5.153s, episode steps: 921, steps per second: 179, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action
137019/1500000: episode: 156, duration: 4.493s, episode steps: 824, steps per second: 183, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
137862/1500000: episode: 157, duration: 4.657s, episode steps: 843, steps per second: 181, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action
138688/1500000: episode: 158, duration: 4.572s, episode steps: 826, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
139526/1500000: episode: 159, duration: 4.653s, episode steps: 838, steps per second: 180, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action
140350/1500000: episode: 160, duration: 4.545s, episode steps: 824, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
141176/1500000: episode: 161, duration: 4.524s, episode steps: 826, steps per second: 183, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
141970/1500000: episode: 162, duration: 4.364s, episode steps: 794, steps per second: 182, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action
142794/1500000: episode: 163, duration: 4.567s, episode steps: 824, steps per second: 180, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
143661/1500000: episode: 164, duration: 4.763s, episode steps: 867, steps per second: 182, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action
144565/1500000: episode: 165, duration: 5.001s, episode steps: 904, steps per second: 181, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action
145390/1500000: episode: 166, duration: 4.676s, episode steps: 825, steps per second: 176, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
146254/1500000: episode: 167, duration: 4.848s, episode steps: 864, steps per second: 178, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action
147078/1500000: episode: 168, duration: 4.625s, episode steps: 824, steps per second: 178, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
147916/1500000: episode: 169, duration: 4.562s, episode steps: 838, steps per second: 184, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action
148740/1500000: episode: 170, duration: 4.572s, episode steps: 824, steps per second: 180, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
149564/1500000: episode: 171, duration: 4.520s, episode steps: 824, steps per second: 182, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
150347/1500000: episode: 172, duration: 4.338s, episode steps: 783, steps per second: 180, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action
151185/1500000: episode: 173, duration: 4.628s, episode steps: 838, steps per second: 181, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action
152095/1500000: episode: 174, duration: 5.071s, episode steps: 910, steps per second: 179, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action
153008/1500000: episode: 175, duration: 5.010s, episode steps: 913, steps per second: 182, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action
153832/1500000: episode: 176, duration: 4.572s, episode steps: 824, steps per second: 180, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
154656/1500000: episode: 177, duration: 4.542s, episode steps: 824, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
155448/1500000: episode: 178, duration: 4.363s, episode steps: 792, steps per second: 182, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action
156272/1500000: episode: 179, duration: 4.591s, episode steps: 824, steps per second: 179, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
157096/1500000: episode: 180, duration: 4.647s, episode steps: 824, steps per second: 177, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
158027/1500000: episode: 181, duration: 5.267s, episode steps: 931, steps per second: 177, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action
158929/1500000: episode: 182, duration: 4.996s, episode steps: 902, steps per second: 181, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action
159753/1500000: episode: 183, duration: 4.614s, episode steps: 824, steps per second: 179, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
160577/1500000: episode: 184, duration: 4.567s, episode steps: 824, steps per second: 180, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
161558/1500000: episode: 185, duration: 5.383s, episode steps: 981, steps per second: 182, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action
162470/1500000: episode: 186, duration: 5.053s, episode steps: 912, steps per second: 180, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action
163373/1500000: episode: 187, duration: 5.002s, episode steps: 903, steps per second: 181, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action
164382/1500000: episode: 188, duration: 5.571s, episode steps: 1009, steps per second: 181, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean actio
165266/1500000: episode: 189, duration: 4.888s, episode steps: 884, steps per second: 181, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action
166257/1500000: episode: 190, duration: 5.429s, episode steps: 991, steps per second: 183, episode reward: -21.000, mean reward: -0.021 [-1.000, 0.000], mean action
167201/1500000: episode: 191, duration: 5.199s, episode steps: 944, steps per second: 182, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action
168114/1500000: episode: 192, duration: 5.181s, episode steps: 913, steps per second: 176, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action
168938/1500000: episode: 193, duration: 4.597s, episode steps: 824, steps per second: 179, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
169762/1500000: episode: 194, duration: 4.555s, episode steps: 824, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
170586/1500000: episode: 195, duration: 4.540s, episode steps: 824, steps per second: 182, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
171410/1500000: episode: 196, duration: 4.506s, episode steps: 824, steps per second: 183, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
172234/1500000: episode: 197, duration: 4.558s, episode steps: 824, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
173060/1500000: episode: 198, duration: 4.564s, episode steps: 826, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
173884/1500000: episode: 199, duration: 4.676s, episode steps: 824, steps per second: 176, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
174857/1500000: episode: 200, duration: 5.391s, episode steps: 973, steps per second: 180, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action
175788/1500000: episode: 201, duration: 5.201s, episode steps: 931, steps per second: 179, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action
176612/1500000: episode: 202, duration: 4.514s, episode steps: 824, steps per second: 183, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
177558/1500000: episode: 203, duration: 5.232s, episode steps: 946, steps per second: 181, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action
178382/1500000: episode: 204, duration: 4.548s, episode steps: 824, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
179364/1500000: episode: 205, duration: 5.596s, episode steps: 982, steps per second: 175, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action
180207/1500000: episode: 206, duration: 4.640s, episode steps: 843, steps per second: 182, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
181031/1500000: episode: 207, duration: 4.545s, episode steps: 824, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
181823/1500000: episode: 208, duration: 4.340s, episode steps: 792, steps per second: 182, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action
182768/1500000: episode: 209, duration: 5.148s, episode steps: 945, steps per second: 184, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action
183714/1500000: episode: 210, duration: 5.212s, episode steps: 946, steps per second: 182, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action
184661/1500000: episode: 211, duration: 5.394s, episode steps: 947, steps per second: 176, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action
185503/1500000: episode: 212, duration: 4.993s, episode steps: 842, steps per second: 169, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
186327/1500000: episode: 213, duration: 4.783s, episode steps: 824, steps per second: 172, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
187186/1500000: episode: 214, duration: 4.929s, episode steps: 859, steps per second: 174, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action
188011/1500000: episode: 215, duration: 4.554s, episode steps: 825, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
188837/1500000: episode: 216, duration: 4.590s, episode steps: 826, steps per second: 180, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
189663/1500000: episode: 217, duration: 4.585s, episode steps: 826, steps per second: 180, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action
```

```
190817/1500000: episode: 218, duration: 6.486s, episode steps: 1154, steps per second: 178, episode reward: -21.000, mean reward: -0.018 [-1.000, 0.000], mean actic
191720/1500000: episode: 219, duration: 5.010s, episode steps: 903, steps per second: 180, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean actic
192544/1500000: episode: 220, duration: 4.546s, episode steps: 824, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
193368/1500000: episode: 221, duration: 4.599s, episode steps: 824, steps per second: 179, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
194192/1500000: episode: 222, duration: 4.591s, episode steps: 824, steps per second: 179, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
194975/1500000: episode: 223, duration: 4.351s, episode steps: 783, steps per second: 180, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean actic
195977/1500000: episode: 224, duration: 5.543s, episode steps: 1002, steps per second: 181, episode reward: -21.000, mean reward: -0.021 [-1.000, 0.000], mean actic
196801/1500000: episode: 225, duration: 4.540s, episode steps: 824, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
197584/1500000: episode: 226, duration: 4.332s, episode steps: 783, steps per second: 181, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean actic
198450/1500000: episode: 227, duration: 4.737s, episode steps: 866, steps per second: 183, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean actic
199616/1500000: episode: 228, duration: 6.394s, episode steps: 1166, steps per second: 182, episode reward: -20.000, mean reward: -0.017 [-1.000, 1.000], mean actic
200497/1500000: episode: 229, duration: 4.892s, episode steps: 881, steps per second: 180, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean actic
201340/1500000: episode: 230, duration: 4.665s, episode steps: 843, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
202164/1500000: episode: 231, duration: 4.609s, episode steps: 824, steps per second: 179, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
203030/1500000: episode: 232, duration: 4.811s, episode steps: 866, steps per second: 180, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean actic
203854/1500000: episode: 233, duration: 4.608s, episode steps: 824, steps per second: 179, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
204678/1500000: episode: 234, duration: 4.549s, episode steps: 824, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
205758/1500000: episode: 235, duration: 5.956s, episode steps: 1080, steps per second: 181, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean actic
206596/1500000: episode: 236, duration: 4.580s, episode steps: 838, steps per second: 183, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean actic
207568/1500000: episode: 237, duration: 5.427s, episode steps: 972, steps per second: 179, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean actic
208515/1500000: episode: 238, duration: 5.173s, episode steps: 947, steps per second: 183, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean actic
209427/1500000: episode: 239, duration: 5.023s, episode steps: 912, steps per second: 182, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean actic
210270/1500000: episode: 240, duration: 4.571s, episode steps: 843, steps per second: 184, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
211094/1500000: episode: 241, duration: 4.563s, episode steps: 824, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
211948/1500000: episode: 242, duration: 4.734s, episode steps: 854, steps per second: 180, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
212878/1500000: episode: 243, duration: 5.158s, episode steps: 930, steps per second: 180, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean actic
213809/1500000: episode: 244, duration: 5.108s, episode steps: 931, steps per second: 182, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean actic
214830/1500000: episode: 245, duration: 5.594s, episode steps: 1021, steps per second: 183, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean actic
215654/1500000: episode: 246, duration: 4.518s, episode steps: 824, steps per second: 182, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
216692/1500000: episode: 247, duration: 5.682s, episode steps: 1038, steps per second: 183, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean actic
217516/1500000: episode: 248, duration: 4.597s, episode steps: 824, steps per second: 179, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
218428/1500000: episode: 249, duration: 5.115s, episode steps: 912, steps per second: 178, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean actic
219418/1500000: episode: 250, duration: 5.427s, episode steps: 990, steps per second: 182, episode reward: -21.000, mean reward: -0.021 [-1.000, 0.000], mean actic
220320/1500000: episode: 251, duration: 4.901s, episode steps: 902, steps per second: 184, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean actic
221144/1500000: episode: 252, duration: 4.491s, episode steps: 824, steps per second: 183, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
222028/1500000: episode: 253, duration: 4.866s, episode steps: 884, steps per second: 182, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean actic
222881/1500000: episode: 254, duration: 4.724s, episode steps: 853, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
224003/1500000: episode: 255, duration: 6.113s, episode steps: 1122, steps per second: 184, episode reward: -21.000, mean reward: -0.019 [-1.000, 0.000], mean actic
224945/1500000: episode: 256, duration: 5.232s, episode steps: 942, steps per second: 180, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean actic
226003/1500000: episode: 257, duration: 5.741s, episode steps: 1058, steps per second: 184, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean actic
226977/1500000: episode: 258, duration: 5.315s, episode steps: 974, steps per second: 183, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean actic
227770/1500000: episode: 259, duration: 4.352s, episode steps: 793, steps per second: 182, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean actic
228687/1500000: episode: 260, duration: 5.031s, episode steps: 917, steps per second: 182, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean actic
229756/1500000: episode: 261, duration: 5.888s, episode steps: 1069, steps per second: 182, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean actic
230705/1500000: episode: 262, duration: 5.216s, episode steps: 949, steps per second: 182, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean actic
231529/1500000: episode: 263, duration: 4.505s, episode steps: 824, steps per second: 183, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
232449/1500000: episode: 264, duration: 5.125s, episode steps: 920, steps per second: 180, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean actic
233381/1500000: episode: 265, duration: 5.142s, episode steps: 932, steps per second: 181, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean actic
234219/1500000: episode: 266, duration: 4.582s, episode steps: 838, steps per second: 183, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean actic
235043/1500000: episode: 267, duration: 4.466s, episode steps: 824, steps per second: 184, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
235895/1500000: episode: 268, duration: 4.700s, episode steps: 852, steps per second: 181, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
236816/1500000: episode: 269, duration: 5.028s, episode steps: 921, steps per second: 183, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean actic
237860/1500000: episode: 270, duration: 5.762s, episode steps: 1044, steps per second: 181, episode reward: -19.000, mean reward: -0.018 [-1.000, 1.000], mean actic
238762/1500000: episode: 271, duration: 5.101s, episode steps: 902, steps per second: 177, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean actic
239741/1500000: episode: 272, duration: 5.378s, episode steps: 979, steps per second: 182, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean actic
240565/1500000: episode: 273, duration: 4.796s, episode steps: 824, steps per second: 172, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
241685/1500000: episode: 274, duration: 6.415s, episode steps: 1120, steps per second: 175, episode reward: -20.000, mean reward: -0.018 [-1.000, 1.000], mean actic
243024/1500000: episode: 275, duration: 7.327s, episode steps: 1339, steps per second: 183, episode reward: -18.000, mean reward: -0.013 [-1.000, 1.000], mean actic
244033/1500000: episode: 276, duration: 5.738s, episode steps: 1009, steps per second: 176, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean actic
245088/1500000: episode: 277, duration: 5.939s, episode steps: 1055, steps per second: 178, episode reward: -20.000, mean reward: -0.020 [-1.000, 0.000], mean actic
245980/1500000: episode: 278, duration: 4.978s, episode steps: 892, steps per second: 179, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean actic
247292/1500000: episode: 279, duration: 7.307s, episode steps: 1312, steps per second: 180, episode reward: -20.000, mean reward: -0.015 [-1.000, 1.000], mean actic
248576/1500000: episode: 280, duration: 7.059s, episode steps: 1284, steps per second: 182, episode reward: -19.000, mean reward: -0.015 [-1.000, 1.000], mean actic
249722/1500000: episode: 281, duration: 6.326s, episode steps: 1146, steps per second: 181, episode reward: -21.000, mean reward: -0.018 [-1.000, 0.000], mean actic
250641/1500000: episode: 282, duration: 5.069s, episode steps: 919, steps per second: 181, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean actic
251621/1500000: episode: 283, duration: 5.411s, episode steps: 980, steps per second: 181, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean actic
252526/1500000: episode: 284, duration: 4.986s, episode steps: 905, steps per second: 182, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean actic
253457/1500000: episode: 285, duration: 5.074s, episode steps: 931, steps per second: 184, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean actic
255082/1500000: episode: 286, duration: 8.951s, episode steps: 1625, steps per second: 182, episode reward: -18.000, mean reward: -0.011 [-1.000, 1.000], mean actic
256236/1500000: episode: 287, duration: 6.355s, episode steps: 1154, steps per second: 182, episode reward: -21.000, mean reward: -0.018 [-1.000, 0.000], mean actic
257416/1500000: episode: 288, duration: 6.404s, episode steps: 1180, steps per second: 184, episode reward: -20.000, mean reward: -0.017 [-1.000, 1.000], mean actic
258424/1500000: episode: 289, duration: 5.585s, episode steps: 1008, steps per second: 180, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean actic
259386/1500000: episode: 290, duration: 5.300s, episode steps: 962, steps per second: 182, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean actic
260576/1500000: episode: 291, duration: 6.533s, episode steps: 1190, steps per second: 182, episode reward: -20.000, mean reward: -0.017 [-1.000, 1.000], mean actic
261784/1500000: episode: 292, duration: 6.652s, episode steps: 1208, steps per second: 182, episode reward: -21.000, mean reward: -0.017 [-1.000, 0.000], mean actic
262931/1500000: episode: 293, duration: 6.173s, episode steps: 1147, steps per second: 186, episode reward: -21.000, mean reward: -0.018 [-1.000, 0.000], mean actic
264215/1500000: episode: 294, duration: 7.036s, episode steps: 1284, steps per second: 182, episode reward: -19.000, mean reward: -0.015 [-1.000, 1.000], mean actic
265357/1500000: episode: 295, duration: 6.286s, episode steps: 1142, steps per second: 182, episode reward: -21.000, mean reward: -0.018 [-1.000, 0.000], mean actic
266655/1500000: episode: 296, duration: 7.057s, episode steps: 1298, steps per second: 184, episode reward: -19.000, mean reward: -0.015 [-1.000, 1.000], mean actic
267860/1500000: episode: 297, duration: 6.594s, episode steps: 1205, steps per second: 183, episode reward: -21.000, mean reward: -0.017 [-1.000, 0.000], mean actic
269067/1500000: episode: 298, duration: 6.553s, episode steps: 1207, steps per second: 184, episode reward: -20.000, mean reward: -0.017 [-1.000, 1.000], mean actic
270413/1500000: episode: 299, duration: 7.355s, episode steps: 1346, steps per second: 183, episode reward: -20.000, mean reward: -0.015 [-1.000, 1.000], mean actic
271718/1500000: episode: 300, duration: 7.082s, episode steps: 1305, steps per second: 184, episode reward: -21.000, mean reward: -0.016 [-1.000, 0.000], mean actic
272934/1500000: episode: 301, duration: 6.661s, episode steps: 1216, steps per second: 183, episode reward: -21.000, mean reward: -0.017 [-1.000, 0.000], mean actic
274036/1500000: episode: 302, duration: 6.030s, episode steps: 1102, steps per second: 183, episode reward: -20.000, mean reward: -0.018 [-1.000, 1.000], mean actic
275435/1500000: episode: 303, duration: 7.656s, episode steps: 1399, steps per second: 183, episode reward: -19.000, mean reward: -0.014 [-1.000, 1.000], mean actic
276807/1500000: episode: 304, duration: 7.545s, episode steps: 1372, steps per second: 182, episode reward: -20.000, mean reward: -0.015 [-1.000, 1.000], mean actic
277779/1500000: episode: 305, duration: 5.288s, episode steps: 972, steps per second: 184, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean actic
278603/1500000: episode: 306, duration: 4.514s, episode steps: 824, steps per second: 183, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean actic
279585/1500000: episode: 307, duration: 5.332s, episode steps: 982, steps per second: 184, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean actic
280949/1500000: episode: 308, duration: 7.433s, episode steps: 1364, steps per second: 183, episode reward: -18.000, mean reward: -0.013 [-1.000, 1.000], mean actic
281991/1500000: episode: 309, duration: 5.747s, episode steps: 1042, steps per second: 181, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean actic
283454/1500000: episode: 310, duration: 8.047s, episode steps: 1463, steps per second: 182, episode reward: -18.000, mean reward: -0.012 [-1.000, 1.000], mean actic
284884/1500000: episode: 311, duration: 7.792s, episode steps: 1430, steps per second: 184, episode reward: -20.000, mean reward: -0.014 [-1.000, 1.000], mean actic
286136/1500000: episode: 312, duration: 6.784s, episode steps: 1252, steps per second: 185, episode reward: -20.000, mean reward: -0.016 [-1.000, 1.000], mean actic
287283/1500000: episode: 313, duration: 6.342s, episode steps: 1147, steps per second: 181, episode reward: -19.000, mean reward: -0.017 [-1.000, 1.000], mean actic
288535/1500000: episode: 314, duration: 6.821s, episode steps: 1252, steps per second: 184, episode reward: -19.000, mean reward: -0.015 [-1.000, 1.000], mean actic
290114/1500000: episode: 315, duration: 8.687s, episode steps: 1579, steps per second: 182, episode reward: -20.000, mean reward: -0.013 [-1.000, 1.000], mean actic
291301/1500000: episode: 316, duration: 6.494s, episode steps: 1187, steps per second: 183, episode reward: -21.000, mean reward: -0.018 [-1.000, 1.000], mean actic
292825/1500000: episode: 317, duration: 8.281s, episode steps: 1524, steps per second: 184, episode reward: -17.000, mean reward: -0.011 [-1.000, 1.000], mean actic
294691/1500000: episode: 318, duration: 10.191s, episode steps: 1866, steps per second: 183, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean acti
296067/1500000: episode: 319, duration: 7.558s, episode steps: 1376, steps per second: 182, episode reward: -19.000, mean reward: -0.014 [-1.000, 1.000], mean actic
297581/1500000: episode: 320, duration: 8.641s, episode steps: 1514, steps per second: 175, episode reward: -18.000, mean reward: -0.012 [-1.000, 1.000], mean actic
299057/1500000: episode: 321, duration: 8.151s, episode steps: 1476, steps per second: 181, episode reward: -19.000, mean reward: -0.013 [-1.000, 1.000], mean actic
300312/1500000: episode: 322, duration: 6.885s, episode steps: 1255, steps per second: 182, episode reward: -19.000, mean reward: -0.015 [-1.000, 1.000], mean actic
301713/1500000: episode: 323, duration: 7.693s, episode steps: 1401, steps per second: 182, episode reward: -18.000, mean reward: -0.013 [-1.000, 1.000], mean actic
303163/1500000: episode: 324, duration: 7.943s, episode steps: 1450, steps per second: 183, episode reward: -19.000, mean reward: -0.013 [-1.000, 1.000], mean actic
304615/1500000: episode: 325, duration: 7.994s, episode steps: 1452, steps per second: 182, episode reward: -18.000, mean reward: -0.012 [-1.000, 1.000], mean actic
306263/1500000: episode: 326, duration: 9.119s, episode steps: 1648, steps per second: 181, episode reward: -17.000, mean reward: -0.010 [-1.000, 1.000], mean actic
```

```
308203/1500000: episode: 327, duration: 10.711s, episode steps: 1940, steps per second: 181, episode reward: -21.000, mean reward: -0.011 [-1.000, 0.000], mean acti
309556/1500000: episode: 328, duration: 7.596s, episode steps: 1353, steps per second: 178, episode reward: -19.000, mean reward: -0.014 [-1.000, 1.000], mean acti
311056/1500000: episode: 329, duration: 8.200s, episode steps: 1500, steps per second: 183, episode reward: -20.000, mean reward: -0.013 [-1.000, 1.000], mean acti
312685/1500000: episode: 330, duration: 8.977s, episode steps: 1629, steps per second: 181, episode reward: -16.000, mean reward: -0.010 [-1.000, 1.000], mean acti
314241/1500000: episode: 331, duration: 8.566s, episode steps: 1556, steps per second: 182, episode reward: -21.000, mean reward: -0.013 [-1.000, 0.000], mean acti
315637/1500000: episode: 332, duration: 7.626s, episode steps: 1396, steps per second: 183, episode reward: -19.000, mean reward: -0.014 [-1.000, 1.000], mean acti
317218/1500000: episode: 333, duration: 8.672s, episode steps: 1581, steps per second: 182, episode reward: -20.000, mean reward: -0.013 [-1.000, 1.000], mean acti
319013/1500000: episode: 334, duration: 9.845s, episode steps: 1795, steps per second: 182, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean acti
320644/1500000: episode: 335, duration: 9.104s, episode steps: 1631, steps per second: 179, episode reward: -19.000, mean reward: -0.012 [-1.000, 1.000], mean acti
322356/1500000: episode: 336, duration: 9.394s, episode steps: 1712, steps per second: 182, episode reward: -17.000, mean reward: -0.010 [-1.000, 1.000], mean acti
323715/1500000: episode: 337, duration: 7.404s, episode steps: 1359, steps per second: 184, episode reward: -20.000, mean reward: -0.015 [-1.000, 1.000], mean acti
325672/1500000: episode: 338, duration: 10.623s, episode steps: 1957, steps per second: 184, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
327633/1500000: episode: 339, duration: 10.737s, episode steps: 1961, steps per second: 183, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean acti
328962/1500000: episode: 340, duration: 7.242s, episode steps: 1329, steps per second: 184, episode reward: -19.000, mean reward: -0.014 [-1.000, 1.000], mean acti
330964/1500000: episode: 341, duration: 10.937s, episode steps: 2002, steps per second: 183, episode reward: -21.000, mean reward: -0.010 [-1.000, 0.000], mean acti
332575/1500000: episode: 342, duration: 8.748s, episode steps: 1611, steps per second: 184, episode reward: -16.000, mean reward: -0.010 [-1.000, 1.000], mean acti
334014/1500000: episode: 343, duration: 7.984s, episode steps: 1439, steps per second: 180, episode reward: -20.000, mean reward: -0.014 [-1.000, 1.000], mean acti
336003/1500000: episode: 344, duration: 10.880s, episode steps: 1989, steps per second: 183, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean acti
338098/1500000: episode: 345, duration: 11.376s, episode steps: 2095, steps per second: 184, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
339735/1500000: episode: 346, duration: 8.984s, episode steps: 1637, steps per second: 182, episode reward: -19.000, mean reward: -0.012 [-1.000, 1.000], mean acti
341418/1500000: episode: 347, duration: 9.151s, episode steps: 1683, steps per second: 184, episode reward: -20.000, mean reward: -0.012 [-1.000, 1.000], mean acti
343144/1500000: episode: 348, duration: 9.438s, episode steps: 1726, steps per second: 183, episode reward: -17.000, mean reward: -0.010 [-1.000, 1.000], mean acti
345355/1500000: episode: 349, duration: 12.124s, episode steps: 2211, steps per second: 182, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
347100/1500000: episode: 350, duration: 9.518s, episode steps: 1745, steps per second: 183, episode reward: -19.000, mean reward: -0.011 [-1.000, 1.000], mean acti
348790/1500000: episode: 351, duration: 9.298s, episode steps: 1690, steps per second: 182, episode reward: -21.000, mean reward: -0.012 [-1.000, 0.000], mean acti
350430/1500000: episode: 352, duration: 9.017s, episode steps: 1640, steps per second: 182, episode reward: -19.000, mean reward: -0.012 [-1.000, 1.000], mean acti
351674/1500000: episode: 353, duration: 6.849s, episode steps: 1244, steps per second: 182, episode reward: -20.000, mean reward: -0.016 [-1.000, 1.000], mean acti
353729/1500000: episode: 354, duration: 11.689s, episode steps: 2055, steps per second: 176, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
355376/1500000: episode: 355, duration: 9.164s, episode steps: 1647, steps per second: 180, episode reward: -20.000, mean reward: -0.012 [-1.000, 1.000], mean acti
356705/1500000: episode: 356, duration: 7.178s, episode steps: 1329, steps per second: 185, episode reward: -19.000, mean reward: -0.014 [-1.000, 1.000], mean acti
358091/1500000: episode: 357, duration: 7.597s, episode steps: 1386, steps per second: 182, episode reward: -21.000, mean reward: -0.015 [-1.000, 0.000], mean acti
359806/1500000: episode: 358, duration: 9.331s, episode steps: 1715, steps per second: 184, episode reward: -17.000, mean reward: -0.010 [-1.000, 1.000], mean acti
361062/1500000: episode: 359, duration: 6.996s, episode steps: 1256, steps per second: 180, episode reward: -21.000, mean reward: -0.017 [-1.000, 0.000], mean acti
362619/1500000: episode: 360, duration: 8.782s, episode steps: 1557, steps per second: 177, episode reward: -20.000, mean reward: -0.013 [-1.000, 1.000], mean acti
364451/1500000: episode: 361, duration: 10.235s, episode steps: 1832, steps per second: 179, episode reward: -20.000, mean reward: -0.011 [-1.000, 1.000], mean acti
366191/1500000: episode: 362, duration: 9.572s, episode steps: 1740, steps per second: 182, episode reward: -20.000, mean reward: -0.011 [-1.000, 1.000], mean acti
367440/1500000: episode: 363, duration: 6.770s, episode steps: 1249, steps per second: 184, episode reward: -20.000, mean reward: -0.016 [-1.000, 1.000], mean acti
368779/1500000: episode: 364, duration: 7.441s, episode steps: 1339, steps per second: 180, episode reward: -20.000, mean reward: -0.015 [-1.000, 1.000], mean acti
370486/1500000: episode: 365, duration: 9.388s, episode steps: 1707, steps per second: 182, episode reward: -18.000, mean reward: -0.011 [-1.000, 1.000], mean acti
372786/1500000: episode: 366, duration: 12.555s, episode steps: 2300, steps per second: 183, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
374769/1500000: episode: 367, duration: 10.931s, episode steps: 1983, steps per second: 181, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean acti
376781/1500000: episode: 368, duration: 10.975s, episode steps: 2012, steps per second: 183, episode reward: -14.000, mean reward: -0.007 [-1.000, 1.000], mean acti
378129/1500000: episode: 369, duration: 7.449s, episode steps: 1348, steps per second: 181, episode reward: -20.000, mean reward: -0.015 [-1.000, 1.000], mean actio
379123/1500000: episode: 370, duration: 5.465s, episode steps: 994, steps per second: 182, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action
380751/1500000: episode: 371, duration: 8.843s, episode steps: 1628, steps per second: 184, episode reward: -16.000, mean reward: -0.010 [-1.000, 1.000], mean acti
382668/1500000: episode: 372, duration: 10.494s, episode steps: 1917, steps per second: 183, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
383736/1500000: episode: 373, duration: 5.876s, episode steps: 1068, steps per second: 182, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean acti
385521/1500000: episode: 374, duration: 9.859s, episode steps: 1785, steps per second: 181, episode reward: -17.000, mean reward: -0.010 [-1.000, 1.000], mean acti
387333/1500000: episode: 375, duration: 9.849s, episode steps: 1812, steps per second: 184, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean acti
389306/1500000: episode: 376, duration: 10.691s, episode steps: 1973, steps per second: 185, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
391105/1500000: episode: 377, duration: 9.831s, episode steps: 1799, steps per second: 183, episode reward: -15.000, mean reward: -0.008 [-1.000, 1.000], mean acti
392436/1500000: episode: 378, duration: 7.295s, episode steps: 1331, steps per second: 182, episode reward: -19.000, mean reward: -0.014 [-1.000, 1.000], mean acti
394236/1500000: episode: 379, duration: 9.860s, episode steps: 1800, steps per second: 183, episode reward: -20.000, mean reward: -0.011 [-1.000, 1.000], mean acti
395923/1500000: episode: 380, duration: 9.267s, episode steps: 1687, steps per second: 182, episode reward: -20.000, mean reward: -0.012 [-1.000, 1.000], mean acti
397071/1500000: episode: 381, duration: 6.353s, episode steps: 1148, steps per second: 181, episode reward: -20.000, mean reward: -0.017 [-1.000, 1.000], mean acti
398857/1500000: episode: 382, duration: 9.809s, episode steps: 1786, steps per second: 182, episode reward: -19.000, mean reward: -0.011 [-1.000, 1.000], mean acti
401001/1500000: episode: 383, duration: 11.653s, episode steps: 2144, steps per second: 184, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
403025/1500000: episode: 384, duration: 11.096s, episode steps: 2024, steps per second: 182, episode reward: -20.000, mean reward: -0.010 [-1.000, 1.000], mean acti
404739/1500000: episode: 385, duration: 9.365s, episode steps: 1714, steps per second: 183, episode reward: -18.000, mean reward: -0.011 [-1.000, 1.000], mean acti
406474/1500000: episode: 386, duration: 9.664s, episode steps: 1735, steps per second: 180, episode reward: -17.000, mean reward: -0.010 [-1.000, 1.000], mean acti
408389/1500000: episode: 387, duration: 10.842s, episode steps: 1915, steps per second: 177, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean acti
410144/1500000: episode: 388, duration: 9.957s, episode steps: 1755, steps per second: 176, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean acti
411918/1500000: episode: 389, duration: 9.835s, episode steps: 1774, steps per second: 180, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean acti
413966/1500000: episode: 390, duration: 11.336s, episode steps: 2048, steps per second: 181, episode reward: -20.000, mean reward: -0.010 [-1.000, 1.000], mean acti
416117/1500000: episode: 391, duration: 11.830s, episode steps: 2151, steps per second: 182, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
417810/1500000: episode: 392, duration: 9.294s, episode steps: 1693, steps per second: 182, episode reward: -21.000, mean reward: -0.012 [-1.000, 0.000], mean acti
419593/1500000: episode: 393, duration: 9.866s, episode steps: 1783, steps per second: 181, episode reward: -19.000, mean reward: -0.011 [-1.000, 1.000], mean acti
421614/1500000: episode: 394, duration: 11.259s, episode steps: 2021, steps per second: 180, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
423027/1500000: episode: 395, duration: 7.775s, episode steps: 1413, steps per second: 182, episode reward: -19.000, mean reward: -0.013 [-1.000, 1.000], mean acti
425347/1500000: episode: 396, duration: 12.710s, episode steps: 2320, steps per second: 183, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
426774/1500000: episode: 397, duration: 7.774s, episode steps: 1427, steps per second: 184, episode reward: -18.000, mean reward: -0.013 [-1.000, 1.000], mean acti
428787/1500000: episode: 398, duration: 10.962s, episode steps: 2013, steps per second: 184, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
430576/1500000: episode: 399, duration: 9.803s, episode steps: 1789, steps per second: 182, episode reward: -16.000, mean reward: -0.009 [-1.000, 1.000], mean acti
432572/1500000: episode: 400, duration: 10.936s, episode steps: 1996, steps per second: 183, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
434043/1500000: episode: 401, duration: 8.060s, episode steps: 1471, steps per second: 183, episode reward: -18.000, mean reward: -0.012 [-1.000, 1.000], mean acti
436075/1500000: episode: 402, duration: 11.159s, episode steps: 2032, steps per second: 182, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
437959/1500000: episode: 403, duration: 10.308s, episode steps: 1884, steps per second: 183, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
439821/1500000: episode: 404, duration: 10.239s, episode steps: 1862, steps per second: 182, episode reward: -20.000, mean reward: -0.011 [-1.000, 1.000], mean acti
442268/1500000: episode: 405, duration: 13.340s, episode steps: 2447, steps per second: 183, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
443732/1500000: episode: 406, duration: 7.919s, episode steps: 1464, steps per second: 185, episode reward: -19.000, mean reward: -0.013 [-1.000, 1.000], mean acti
445333/1500000: episode: 407, duration: 8.709s, episode steps: 1601, steps per second: 184, episode reward: -19.000, mean reward: -0.012 [-1.000, 1.000], mean acti
447203/1500000: episode: 408, duration: 10.159s, episode steps: 1870, steps per second: 184, episode reward: -21.000, mean reward: -0.011 [-1.000, 0.000], mean acti
449020/1500000: episode: 409, duration: 9.914s, episode steps: 1817, steps per second: 183, episode reward: -16.000, mean reward: -0.009 [-1.000, 1.000], mean acti
451115/1500000: episode: 410, duration: 11.573s, episode steps: 2095, steps per second: 181, episode reward: -20.000, mean reward: -0.010 [-1.000, 1.000], mean acti
453248/1500000: episode: 411, duration: 11.655s, episode steps: 2133, steps per second: 183, episode reward: -19.000, mean reward: -0.009 [-1.000, 1.000], mean acti
455463/1500000: episode: 412, duration: 12.085s, episode steps: 2215, steps per second: 183, episode reward: -20.000, mean reward: -0.009 [-1.000, 1.000], mean acti
457511/1500000: episode: 413, duration: 11.348s, episode steps: 2048, steps per second: 180, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
459366/1500000: episode: 414, duration: 10.138s, episode steps: 1855, steps per second: 183, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean acti
461029/1500000: episode: 415, duration: 9.053s, episode steps: 1663, steps per second: 184, episode reward: -18.000, mean reward: -0.011 [-1.000, 1.000], mean acti
463141/1500000: episode: 416, duration: 11.795s, episode steps: 2112, steps per second: 179, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
464788/1500000: episode: 417, duration: 9.276s, episode steps: 1647, steps per second: 178, episode reward: -18.000, mean reward: -0.011 [-1.000, 1.000], mean acti
466844/1500000: episode: 418, duration: 11.621s, episode steps: 2056, steps per second: 177, episode reward: -19.000, mean reward: -0.009 [-1.000, 1.000], mean acti
468820/1500000: episode: 419, duration: 10.814s, episode steps: 1976, steps per second: 183, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean acti
470598/1500000: episode: 420, duration: 9.777s, episode steps: 1778, steps per second: 182, episode reward: -19.000, mean reward: -0.011 [-1.000, 1.000], mean acti
472237/1500000: episode: 421, duration: 8.932s, episode steps: 1639, steps per second: 183, episode reward: -21.000, mean reward: -0.013 [-1.000, 0.000], mean actio
473952/1500000: episode: 422, duration: 9.350s, episode steps: 1715, steps per second: 183, episode reward: -20.000, mean reward: -0.012 [-1.000, 1.000], mean acti
475924/1500000: episode: 423, duration: 10.762s, episode steps: 1972, steps per second: 183, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
477767/1500000: episode: 424, duration: 10.093s, episode steps: 1843, steps per second: 183, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean acti
479851/1500000: episode: 425, duration: 11.489s, episode steps: 2084, steps per second: 181, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
481637/1500000: episode: 426, duration: 9.758s, episode steps: 1786, steps per second: 183, episode reward: -21.000, mean reward: -0.012 [-1.000, 0.000], mean actio
483786/1500000: episode: 427, duration: 11.816s, episode steps: 2149, steps per second: 182, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
486103/1500000: episode: 428, duration: 12.607s, episode steps: 2317, steps per second: 184, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
488375/1500000: episode: 429, duration: 12.366s, episode steps: 2272, steps per second: 184, episode reward: -19.000, mean reward: -0.008 [-1.000, 1.000], mean acti
490666/1500000: episode: 430, duration: 12.377s, episode steps: 2291, steps per second: 185, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
492894/1500000: episode: 431, duration: 12.061s, episode steps: 2228, steps per second: 185, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
494753/1500000: episode: 432, duration: 10.336s, episode steps: 1859, steps per second: 180, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean acti
496344/1500000: episode: 433, duration: 8.693s, episode steps: 1591, steps per second: 183, episode reward: -19.000, mean reward: -0.012 [-1.000, 1.000], mean acti
498484/1500000: episode: 434, duration: 11.626s, episode steps: 2140, steps per second: 184, episode reward: -13.000, mean reward: -0.006 [-1.000, 1.000], mean acti
499930/1500000: episode: 435, duration: 7.898s, episode steps: 1446, steps per second: 183, episode reward: -20.000, mean reward: -0.014 [-1.000, 1.000], mean actio
```

```
501545/1500000: episode: 436, duration: 8.756s, episode steps: 1615, steps per second: 184, episode reward: -19.000, mean reward: -0.012 [-1.000, 1.000], mean actic
503002/1500000: episode: 437, duration: 7.972s, episode steps: 1457, steps per second: 183, episode reward: -21.000, mean reward: -0.014 [-1.000, 0.000], mean actic
504907/1500000: episode: 438, duration: 10.421s, episode steps: 1905, steps per second: 183, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
506728/1500000: episode: 439, duration: 9.931s, episode steps: 1821, steps per second: 183, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean acti
509019/1500000: episode: 440, duration: 12.491s, episode steps: 2291, steps per second: 183, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
511040/1500000: episode: 441, duration: 11.055s, episode steps: 2021, steps per second: 183, episode reward: -20.000, mean reward: -0.010 [-1.000, 1.000], mean acti
513054/1500000: episode: 442, duration: 10.989s, episode steps: 2014, steps per second: 183, episode reward: -19.000, mean reward: -0.009 [-1.000, 1.000], mean acti
515240/1500000: episode: 443, duration: 11.828s, episode steps: 2186, steps per second: 185, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
517185/1500000: episode: 444, duration: 10.679s, episode steps: 1945, steps per second: 182, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean acti
519004/1500000: episode: 445, duration: 10.035s, episode steps: 1819, steps per second: 181, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean acti
520860/1500000: episode: 446, duration: 10.014s, episode steps: 1856, steps per second: 185, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean acti
522890/1500000: episode: 447, duration: 11.507s, episode steps: 2030, steps per second: 176, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
524585/1500000: episode: 448, duration: 9.461s, episode steps: 1695, steps per second: 179, episode reward: -18.000, mean reward: -0.011 [-1.000, 1.000], mean actic
526430/1500000: episode: 449, duration: 10.124s, episode steps: 1845, steps per second: 182, episode reward: -21.000, mean reward: -0.011 [-1.000, 0.000], mean acti
528775/1500000: episode: 450, duration: 12.908s, episode steps: 2345, steps per second: 182, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
530420/1500000: episode: 451, duration: 9.187s, episode steps: 1645, steps per second: 179, episode reward: -21.000, mean reward: -0.013 [-1.000, 0.000], mean actic
532297/1500000: episode: 452, duration: 10.240s, episode steps: 1877, steps per second: 183, episode reward: -21.000, mean reward: -0.011 [-1.000, 0.000], mean acti
534573/1500000: episode: 453, duration: 12.445s, episode steps: 2276, steps per second: 183, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
536418/1500000: episode: 454, duration: 10.124s, episode steps: 1845, steps per second: 182, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
537907/1500000: episode: 455, duration: 8.261s, episode steps: 1489, steps per second: 180, episode reward: -18.000, mean reward: -0.012 [-1.000, 1.000], mean actic
539924/1500000: episode: 456, duration: 11.199s, episode steps: 2017, steps per second: 180, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
542166/1500000: episode: 457, duration: 12.179s, episode steps: 2242, steps per second: 184, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
544091/1500000: episode: 458, duration: 10.532s, episode steps: 1925, steps per second: 183, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
546303/1500000: episode: 459, duration: 12.033s, episode steps: 2212, steps per second: 184, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
548366/1500000: episode: 460, duration: 11.208s, episode steps: 2063, steps per second: 184, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
550165/1500000: episode: 461, duration: 9.846s, episode steps: 1799, steps per second: 183, episode reward: -19.000, mean reward: -0.011 [-1.000, 1.000], mean acti
551888/1500000: episode: 462, duration: 9.420s, episode steps: 1723, steps per second: 183, episode reward: -17.000, mean reward: -0.010 [-1.000, 1.000], mean actic
553782/1500000: episode: 463, duration: 10.450s, episode steps: 1894, steps per second: 181, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean acti
555718/1500000: episode: 464, duration: 10.529s, episode steps: 1936, steps per second: 184, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
557777/1500000: episode: 465, duration: 11.263s, episode steps: 2059, steps per second: 183, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
560472/1500000: episode: 466, duration: 14.764s, episode steps: 2695, steps per second: 183, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
562739/1500000: episode: 467, duration: 12.376s, episode steps: 2267, steps per second: 183, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
565093/1500000: episode: 468, duration: 12.800s, episode steps: 2354, steps per second: 184, episode reward: -19.000, mean reward: -0.008 [-1.000, 1.000], mean acti
567104/1500000: episode: 469, duration: 11.142s, episode steps: 2011, steps per second: 180, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
569125/1500000: episode: 470, duration: 11.093s, episode steps: 2021, steps per second: 182, episode reward: -19.000, mean reward: -0.009 [-1.000, 1.000], mean acti
570706/1500000: episode: 471, duration: 8.702s, episode steps: 1581, steps per second: 182, episode reward: -20.000, mean reward: -0.013 [-1.000, 1.000], mean acti
572494/1500000: episode: 472, duration: 9.817s, episode steps: 1788, steps per second: 182, episode reward: -19.000, mean reward: -0.011 [-1.000, 1.000], mean actic
574215/1500000: episode: 473, duration: 9.545s, episode steps: 1721, steps per second: 180, episode reward: -20.000, mean reward: -0.012 [-1.000, 1.000], mean acti
576015/1500000: episode: 474, duration: 9.812s, episode steps: 1800, steps per second: 183, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean acti
578080/1500000: episode: 475, duration: 11.383s, episode steps: 2065, steps per second: 181, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
579888/1500000: episode: 476, duration: 10.252s, episode steps: 1808, steps per second: 176, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
581694/1500000: episode: 477, duration: 9.994s, episode steps: 1806, steps per second: 181, episode reward: -20.000, mean reward: -0.011 [-1.000, 1.000], mean acti
584407/1500000: episode: 478, duration: 14.797s, episode steps: 2713, steps per second: 183, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
586421/1500000: episode: 479, duration: 11.026s, episode steps: 2014, steps per second: 183, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
588061/1500000: episode: 480, duration: 8.917s, episode steps: 1640, steps per second: 184, episode reward: -19.000, mean reward: -0.012 [-1.000, 1.000], mean acti
590416/1500000: episode: 481, duration: 12.861s, episode steps: 2355, steps per second: 183, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
591910/1500000: episode: 482, duration: 8.253s, episode steps: 1494, steps per second: 181, episode reward: -19.000, mean reward: -0.013 [-1.000, 1.000], mean actic
593782/1500000: episode: 483, duration: 10.292s, episode steps: 1872, steps per second: 182, episode reward: -16.000, mean reward: -0.009 [-1.000, 1.000], mean acti
596396/1500000: episode: 484, duration: 14.330s, episode steps: 2614, steps per second: 182, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
598282/1500000: episode: 485, duration: 10.333s, episode steps: 1886, steps per second: 183, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
600450/1500000: episode: 486, duration: 11.854s, episode steps: 2168, steps per second: 183, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
602267/1500000: episode: 487, duration: 10.031s, episode steps: 1817, steps per second: 181, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean acti
604592/1500000: episode: 488, duration: 12.798s, episode steps: 2325, steps per second: 182, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
607610/1500000: episode: 489, duration: 16.545s, episode steps: 3018, steps per second: 182, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
610056/1500000: episode: 490, duration: 13.576s, episode steps: 2446, steps per second: 180, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
612340/1500000: episode: 491, duration: 12.629s, episode steps: 2284, steps per second: 181, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
614498/1500000: episode: 492, duration: 11.964s, episode steps: 2158, steps per second: 180, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
616809/1500000: episode: 493, duration: 12.613s, episode steps: 2311, steps per second: 183, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
619084/1500000: episode: 494, duration: 12.400s, episode steps: 2275, steps per second: 183, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
620963/1500000: episode: 495, duration: 10.305s, episode steps: 1879, steps per second: 182, episode reward: -21.000, mean reward: -0.011 [-1.000, 0.000], mean acti
623116/1500000: episode: 496, duration: 11.779s, episode steps: 2153, steps per second: 183, episode reward: -19.000, mean reward: -0.009 [-1.000, 1.000], mean acti
624688/1500000: episode: 497, duration: 8.605s, episode steps: 1572, steps per second: 183, episode reward: -19.000, mean reward: -0.012 [-1.000, 1.000], mean actic
627485/1500000: episode: 498, duration: 15.440s, episode steps: 2797, steps per second: 181, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
629699/1500000: episode: 499, duration: 12.240s, episode steps: 2214, steps per second: 181, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
631866/1500000: episode: 500, duration: 11.826s, episode steps: 2167, steps per second: 183, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
634296/1500000: episode: 501, duration: 13.412s, episode steps: 2430, steps per second: 181, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
636145/1500000: episode: 502, duration: 10.526s, episode steps: 1849, steps per second: 176, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean acti
637883/1500000: episode: 503, duration: 9.506s, episode steps: 1738, steps per second: 183, episode reward: -17.000, mean reward: -0.010 [-1.000, 1.000], mean acti
640073/1500000: episode: 504, duration: 11.959s, episode steps: 2190, steps per second: 183, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
641786/1500000: episode: 505, duration: 9.430s, episode steps: 1713, steps per second: 182, episode reward: -17.000, mean reward: -0.010 [-1.000, 1.000], mean actic
643551/1500000: episode: 506, duration: 9.585s, episode steps: 1765, steps per second: 184, episode reward: -19.000, mean reward: -0.011 [-1.000, 1.000], mean acti
645507/1500000: episode: 507, duration: 10.680s, episode steps: 1956, steps per second: 183, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
647467/1500000: episode: 508, duration: 10.791s, episode steps: 1960, steps per second: 182, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
649382/1500000: episode: 509, duration: 10.477s, episode steps: 1915, steps per second: 183, episode reward: -20.000, mean reward: -0.010 [-1.000, 1.000], mean acti
651084/1500000: episode: 510, duration: 9.270s, episode steps: 1702, steps per second: 184, episode reward: -19.000, mean reward: -0.011 [-1.000, 1.000], mean acti
653775/1500000: episode: 511, duration: 14.659s, episode steps: 2691, steps per second: 184, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
655502/1500000: episode: 512, duration: 9.523s, episode steps: 1727, steps per second: 181, episode reward: -20.000, mean reward: -0.012 [-1.000, 1.000], mean acti
657607/1500000: episode: 513, duration: 11.647s, episode steps: 2105, steps per second: 181, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
659984/1500000: episode: 514, duration: 12.999s, episode steps: 2377, steps per second: 183, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
662076/1500000: episode: 515, duration: 11.393s, episode steps: 2092, steps per second: 184, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
664604/1500000: episode: 516, duration: 14.009s, episode steps: 2528, steps per second: 180, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
666850/1500000: episode: 517, duration: 12.297s, episode steps: 2246, steps per second: 183, episode reward: -19.000, mean reward: -0.008 [-1.000, 1.000], mean acti
668884/1500000: episode: 518, duration: 11.370s, episode steps: 2034, steps per second: 179, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
671272/1500000: episode: 519, duration: 13.259s, episode steps: 2388, steps per second: 180, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
673827/1500000: episode: 520, duration: 13.933s, episode steps: 2555, steps per second: 183, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
675705/1500000: episode: 521, duration: 10.342s, episode steps: 1878, steps per second: 182, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean acti
677645/1500000: episode: 522, duration: 10.568s, episode steps: 1940, steps per second: 184, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean acti
679460/1500000: episode: 523, duration: 9.989s, episode steps: 1815, steps per second: 182, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean actic
681291/1500000: episode: 524, duration: 10.006s, episode steps: 1831, steps per second: 183, episode reward: -21.000, mean reward: -0.011 [-1.000, 0.000], mean acti
683802/1500000: episode: 525, duration: 13.725s, episode steps: 2511, steps per second: 183, episode reward: -19.000, mean reward: -0.008 [-1.000, 1.000], mean acti
686212/1500000: episode: 526, duration: 13.318s, episode steps: 2410, steps per second: 181, episode reward: -18.000, mean reward: -0.007 [-1.000, 1.000], mean acti
688577/1500000: episode: 527, duration: 12.940s, episode steps: 2365, steps per second: 183, episode reward: -20.000, mean reward: -0.008 [-1.000, 1.000], mean acti
690994/1500000: episode: 528, duration: 13.358s, episode steps: 2417, steps per second: 181, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
692737/1500000: episode: 529, duration: 9.718s, episode steps: 1743, steps per second: 179, episode reward: -19.000, mean reward: -0.011 [-1.000, 1.000], mean actic
695162/1500000: episode: 530, duration: 13.158s, episode steps: 2425, steps per second: 184, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
697330/1500000: episode: 531, duration: 11.698s, episode steps: 2168, steps per second: 185, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
699417/1500000: episode: 532, duration: 11.475s, episode steps: 2087, steps per second: 182, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
701662/1500000: episode: 533, duration: 12.221s, episode steps: 2245, steps per second: 184, episode reward: -19.000, mean reward: -0.008 [-1.000, 1.000], mean acti
703828/1500000: episode: 534, duration: 11.633s, episode steps: 2166, steps per second: 186, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
706046/1500000: episode: 535, duration: 12.099s, episode steps: 2218, steps per second: 183, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
708551/1500000: episode: 536, duration: 13.486s, episode steps: 2505, steps per second: 186, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
710676/1500000: episode: 537, duration: 11.505s, episode steps: 2125, steps per second: 185, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
713335/1500000: episode: 538, duration: 14.540s, episode steps: 2659, steps per second: 183, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
715512/1500000: episode: 539, duration: 11.970s, episode steps: 2177, steps per second: 182, episode reward: -20.000, mean reward: -0.009 [-1.000, 1.000], mean acti
717496/1500000: episode: 540, duration: 10.931s, episode steps: 1984, steps per second: 182, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
719833/1500000: episode: 541, duration: 12.726s, episode steps: 2337, steps per second: 184, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
722260/1500000: episode: 542, duration: 13.596s, episode steps: 2427, steps per second: 179, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
724604/1500000: episode: 543, duration: 12.871s, episode steps: 2344, steps per second: 182, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
726971/1500000: episode: 544, duration: 12.940s, episode steps: 2367, steps per second: 183, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
```

```
729067/1500000: episode: 545, duration: 11.798s, episode steps: 2096, steps per second: 178, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
731185/1500000: episode: 546, duration: 11.546s, episode steps: 2118, steps per second: 183, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
733545/1500000: episode: 547, duration: 12.889s, episode steps: 2360, steps per second: 183, episode reward: -19.000, mean reward: -0.008 [-1.000, 1.000], mean acti
735877/1500000: episode: 548, duration: 12.811s, episode steps: 2332, steps per second: 182, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
737966/1500000: episode: 549, duration: 11.273s, episode steps: 2089, steps per second: 185, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
740123/1500000: episode: 550, duration: 11.863s, episode steps: 2157, steps per second: 182, episode reward: -20.000, mean reward: -0.009 [-1.000, 1.000], mean acti
742960/1500000: episode: 551, duration: 15.634s, episode steps: 2837, steps per second: 181, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
745120/1500000: episode: 552, duration: 11.884s, episode steps: 2160, steps per second: 182, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
747465/1500000: episode: 553, duration: 13.008s, episode steps: 2345, steps per second: 180, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
749928/1500000: episode: 554, duration: 13.515s, episode steps: 2463, steps per second: 182, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
752907/1500000: episode: 555, duration: 16.175s, episode steps: 2979, steps per second: 184, episode reward: -17.000, mean reward: -0.006 [-1.000, 1.000], mean acti
755008/1500000: episode: 556, duration: 11.432s, episode steps: 2101, steps per second: 184, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
757635/1500000: episode: 557, duration: 14.447s, episode steps: 2627, steps per second: 182, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
759959/1500000: episode: 558, duration: 12.594s, episode steps: 2324, steps per second: 185, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
762642/1500000: episode: 559, duration: 14.563s, episode steps: 2683, steps per second: 184, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
765305/1500000: episode: 560, duration: 14.471s, episode steps: 2663, steps per second: 184, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
767615/1500000: episode: 561, duration: 12.775s, episode steps: 2310, steps per second: 181, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
770068/1500000: episode: 562, duration: 13.387s, episode steps: 2453, steps per second: 183, episode reward: -20.000, mean reward: -0.008 [-1.000, 1.000], mean acti
772748/1500000: episode: 563, duration: 14.680s, episode steps: 2680, steps per second: 183, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
774677/1500000: episode: 564, duration: 10.918s, episode steps: 1929, steps per second: 177, episode reward: -15.000, mean reward: -0.008 [-1.000, 1.000], mean acti
776892/1500000: episode: 565, duration: 12.135s, episode steps: 2215, steps per second: 183, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
779021/1500000: episode: 566, duration: 11.659s, episode steps: 2129, steps per second: 183, episode reward: -14.000, mean reward: -0.007 [-1.000, 1.000], mean acti
781124/1500000: episode: 567, duration: 11.501s, episode steps: 2103, steps per second: 183, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
783287/1500000: episode: 568, duration: 11.776s, episode steps: 2163, steps per second: 184, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
785790/1500000: episode: 569, duration: 13.734s, episode steps: 2503, steps per second: 182, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
788371/1500000: episode: 570, duration: 14.204s, episode steps: 2581, steps per second: 182, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
790718/1500000: episode: 571, duration: 12.729s, episode steps: 2347, steps per second: 184, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
792736/1500000: episode: 572, duration: 10.975s, episode steps: 2018, steps per second: 184, episode reward: -19.000, mean reward: -0.009 [-1.000, 1.000], mean acti
794801/1500000: episode: 573, duration: 11.257s, episode steps: 2065, steps per second: 183, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
797376/1500000: episode: 574, duration: 14.212s, episode steps: 2575, steps per second: 181, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
799686/1500000: episode: 575, duration: 12.650s, episode steps: 2310, steps per second: 183, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
801867/1500000: episode: 576, duration: 11.966s, episode steps: 2181, steps per second: 182, episode reward: -19.000, mean reward: -0.009 [-1.000, 1.000], mean acti
805002/1500000: episode: 577, duration: 17.393s, episode steps: 3135, steps per second: 180, episode reward: -16.000, mean reward: -0.005 [-1.000, 1.000], mean acti
807358/1500000: episode: 578, duration: 12.698s, episode steps: 2356, steps per second: 186, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
809830/1500000: episode: 579, duration: 13.432s, episode steps: 2472, steps per second: 184, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
811977/1500000: episode: 580, duration: 11.665s, episode steps: 2147, steps per second: 184, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
814134/1500000: episode: 581, duration: 11.605s, episode steps: 2157, steps per second: 186, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
816340/1500000: episode: 582, duration: 12.012s, episode steps: 2206, steps per second: 184, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
818913/1500000: episode: 583, duration: 13.897s, episode steps: 2573, steps per second: 185, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
820949/1500000: episode: 584, duration: 11.038s, episode steps: 2036, steps per second: 184, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
822791/1500000: episode: 585, duration: 9.974s, episode steps: 1842, steps per second: 185, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean actic
824868/1500000: episode: 586, duration: 11.350s, episode steps: 2077, steps per second: 183, episode reward: -19.000, mean reward: -0.009 [-1.000, 1.000], mean acti
827285/1500000: episode: 587, duration: 13.083s, episode steps: 2417, steps per second: 185, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
829843/1500000: episode: 588, duration: 13.984s, episode steps: 2558, steps per second: 183, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
832162/1500000: episode: 589, duration: 12.749s, episode steps: 2319, steps per second: 182, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
834090/1500000: episode: 590, duration: 10.453s, episode steps: 1928, steps per second: 184, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
837449/1500000: episode: 591, duration: 18.551s, episode steps: 3359, steps per second: 181, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
840349/1500000: episode: 592, duration: 15.735s, episode steps: 2900, steps per second: 184, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
843362/1500000: episode: 593, duration: 16.599s, episode steps: 3013, steps per second: 182, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
845980/1500000: episode: 594, duration: 14.448s, episode steps: 2618, steps per second: 181, episode reward: -16.000, mean reward: -0.006 [-1.000, 1.000], mean acti
848585/1500000: episode: 595, duration: 14.291s, episode steps: 2605, steps per second: 182, episode reward: -18.000, mean reward: -0.007 [-1.000, 1.000], mean acti
851325/1500000: episode: 596, duration: 15.140s, episode steps: 2740, steps per second: 181, episode reward: -16.000, mean reward: -0.006 [-1.000, 1.000], mean acti
854468/1500000: episode: 597, duration: 17.235s, episode steps: 3143, steps per second: 182, episode reward: -14.000, mean reward: -0.004 [-1.000, 1.000], mean acti
856879/1500000: episode: 598, duration: 13.111s, episode steps: 2411, steps per second: 184, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
859477/1500000: episode: 599, duration: 14.168s, episode steps: 2598, steps per second: 183, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
861595/1500000: episode: 600, duration: 11.960s, episode steps: 2118, steps per second: 177, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
863876/1500000: episode: 601, duration: 12.343s, episode steps: 2281, steps per second: 185, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
865725/1500000: episode: 602, duration: 10.077s, episode steps: 1849, steps per second: 183, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean acti
868282/1500000: episode: 603, duration: 13.779s, episode steps: 2557, steps per second: 186, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
870234/1500000: episode: 604, duration: 10.573s, episode steps: 1952, steps per second: 185, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
872565/1500000: episode: 605, duration: 12.617s, episode steps: 2331, steps per second: 185, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
874480/1500000: episode: 606, duration: 10.494s, episode steps: 1915, steps per second: 182, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
876382/1500000: episode: 607, duration: 10.385s, episode steps: 1902, steps per second: 183, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
878248/1500000: episode: 608, duration: 10.191s, episode steps: 1866, steps per second: 183, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
880574/1500000: episode: 609, duration: 12.558s, episode steps: 2326, steps per second: 185, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
882729/1500000: episode: 610, duration: 11.774s, episode steps: 2155, steps per second: 183, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
884927/1500000: episode: 611, duration: 11.904s, episode steps: 2198, steps per second: 185, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
887083/1500000: episode: 612, duration: 11.828s, episode steps: 2156, steps per second: 182, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
889626/1500000: episode: 613, duration: 13.983s, episode steps: 2543, steps per second: 182, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
891611/1500000: episode: 614, duration: 10.804s, episode steps: 1985, steps per second: 184, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
893270/1500000: episode: 615, duration: 9.106s, episode steps: 1659, steps per second: 182, episode reward: -18.000, mean reward: -0.011 [-1.000, 1.000], mean actic
895761/1500000: episode: 616, duration: 13.531s, episode steps: 2491, steps per second: 184, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
898326/1500000: episode: 617, duration: 14.403s, episode steps: 2565, steps per second: 178, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
901664/1500000: episode: 618, duration: 18.815s, episode steps: 3338, steps per second: 177, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
903946/1500000: episode: 619, duration: 12.531s, episode steps: 2282, steps per second: 182, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
906110/1500000: episode: 620, duration: 11.911s, episode steps: 2164, steps per second: 182, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
908677/1500000: episode: 621, duration: 14.217s, episode steps: 2567, steps per second: 181, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean acti
911133/1500000: episode: 622, duration: 13.412s, episode steps: 2456, steps per second: 183, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
913778/1500000: episode: 623, duration: 14.508s, episode steps: 2645, steps per second: 182, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
915732/1500000: episode: 624, duration: 10.664s, episode steps: 1954, steps per second: 183, episode reward: -15.000, mean reward: -0.008 [-1.000, 1.000], mean acti
918069/1500000: episode: 625, duration: 13.383s, episode steps: 2337, steps per second: 175, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
920079/1500000: episode: 626, duration: 11.151s, episode steps: 2010, steps per second: 180, episode reward: -14.000, mean reward: -0.007 [-1.000, 1.000], mean acti
922853/1500000: episode: 627, duration: 15.237s, episode steps: 2774, steps per second: 182, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean actic
925023/1500000: episode: 628, duration: 12.001s, episode steps: 2170, steps per second: 181, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
928327/1500000: episode: 629, duration: 18.102s, episode steps: 3304, steps per second: 183, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
930952/1500000: episode: 630, duration: 14.397s, episode steps: 2625, steps per second: 182, episode reward: -16.000, mean reward: -0.006 [-1.000, 1.000], mean acti
933420/1500000: episode: 631, duration: 13.632s, episode steps: 2468, steps per second: 181, episode reward: -18.000, mean reward: -0.007 [-1.000, 1.000], mean acti
935647/1500000: episode: 632, duration: 12.175s, episode steps: 2227, steps per second: 183, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
937501/1500000: episode: 633, duration: 10.065s, episode steps: 1854, steps per second: 184, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean acti
940309/1500000: episode: 634, duration: 15.364s, episode steps: 2808, steps per second: 183, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
942102/1500000: episode: 635, duration: 9.807s, episode steps: 1793, steps per second: 183, episode reward: -20.000, mean reward: -0.011 [-1.000, 1.000], mean actic
943772/1500000: episode: 636, duration: 9.111s, episode steps: 1670, steps per second: 183, episode reward: -19.000, mean reward: -0.011 [-1.000, 1.000], mean actic
945750/1500000: episode: 637, duration: 10.818s, episode steps: 1978, steps per second: 183, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
948428/1500000: episode: 638, duration: 14.671s, episode steps: 2678, steps per second: 183, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
949782/1500000: episode: 639, duration: 7.393s, episode steps: 1354, steps per second: 183, episode reward: -19.000, mean reward: -0.014 [-1.000, 1.000], mean actic
951283/1500000: episode: 640, duration: 8.259s, episode steps: 1501, steps per second: 182, episode reward: -20.000, mean reward: -0.013 [-1.000, 1.000], mean actic
953339/1500000: episode: 641, duration: 11.249s, episode steps: 2056, steps per second: 183, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
955416/1500000: episode: 642, duration: 11.269s, episode steps: 2077, steps per second: 184, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
957974/1500000: episode: 643, duration: 13.960s, episode steps: 2558, steps per second: 183, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean acti
960196/1500000: episode: 644, duration: 12.016s, episode steps: 2222, steps per second: 185, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
962484/1500000: episode: 645, duration: 12.803s, episode steps: 2288, steps per second: 179, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
964754/1500000: episode: 646, duration: 12.425s, episode steps: 2270, steps per second: 183, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
967487/1500000: episode: 647, duration: 14.890s, episode steps: 2733, steps per second: 184, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean acti
970198/1500000: episode: 648, duration: 14.808s, episode steps: 2711, steps per second: 183, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
972157/1500000: episode: 649, duration: 10.698s, episode steps: 1959, steps per second: 183, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
975209/1500000: episode: 650, duration: 17.118s, episode steps: 3052, steps per second: 178, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean actic
977389/1500000: episode: 651, duration: 12.099s, episode steps: 2180, steps per second: 180, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
979506/1500000: episode: 652, duration: 11.593s, episode steps: 2117, steps per second: 183, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean acti
981079/1500000: episode: 653, duration: 8.571s, episode steps: 1573, steps per second: 184, episode reward: -18.000, mean reward: -0.011 [-1.000, 1.000], mean actic
```

```
 983513/1500000: episode:  654, duration: 13.357s, episode steps: 2434, steps per second: 182, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean act
 985769/1500000: episode:  655, duration: 12.451s, episode steps: 2256, steps per second: 181, episode reward: -13.000, mean reward: -0.006 [-1.000, 1.000], mean act
 987936/1500000: episode:  656, duration: 11.800s, episode steps: 2167, steps per second: 184, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean act
 990661/1500000: episode:  657, duration: 14.875s, episode steps: 2725, steps per second: 183, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean act
 992914/1500000: episode:  658, duration: 12.183s, episode steps: 2253, steps per second: 185, episode reward: -13.000, mean reward: -0.006 [-1.000, 1.000], mean act
 995319/1500000: episode:  659, duration: 13.068s, episode steps: 2405, steps per second: 184, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean act
 997653/1500000: episode:  660, duration: 12.694s, episode steps: 2334, steps per second: 184, episode reward: -11.000, mean reward: -0.005 [-1.000, 1.000], mean act
 999601/1500000: episode:  661, duration: 10.546s, episode steps: 1948, steps per second: 185, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean act
1002414/1500000: episode:  662, duration: 15.302s, episode steps: 2813, steps per second: 184, episode reward: -16.000, mean reward: -0.006 [-1.000, 1.000], mean act
1004880/1500000: episode:  663, duration: 13.367s, episode steps: 2466, steps per second: 184, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean act
1007275/1500000: episode:  664, duration: 13.031s, episode steps: 2395, steps per second: 184, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean act
1009399/1500000: episode:  665, duration: 11.448s, episode steps: 2124, steps per second: 186, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean act
1011681/1500000: episode:  666, duration: 12.259s, episode steps: 2282, steps per second: 186, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean act
1014085/1500000: episode:  667, duration: 12.924s, episode steps: 2404, steps per second: 186, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean act
1016277/1500000: episode:  668, duration: 11.783s, episode steps: 2192, steps per second: 186, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean act
1019211/1500000: episode:  669, duration: 15.992s, episode steps: 2934, steps per second: 183, episode reward:  -4.000, mean reward: -0.001 [-1.000, 1.000], mean acti
1021334/1500000: episode:  670, duration: 11.594s, episode steps: 2123, steps per second: 183, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean act
1023369/1500000: episode:  671, duration: 11.027s, episode steps: 2035, steps per second: 185, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean act
1026012/1500000: episode:  672, duration: 14.376s, episode steps: 2643, steps per second: 184, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean act
1028466/1500000: episode:  673, duration: 13.209s, episode steps: 2454, steps per second: 186, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean act
1031071/1500000: episode:  674, duration: 14.484s, episode steps: 2605, steps per second: 180, episode reward:  -8.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1033284/1500000: episode:  675, duration: 11.882s, episode steps: 2213, steps per second: 186, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean act
1035333/1500000: episode:  676, duration: 11.172s, episode steps: 2049, steps per second: 183, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean act
1037316/1500000: episode:  677, duration: 10.727s, episode steps: 1983, steps per second: 184, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean act
1039055/1500000: episode:  678, duration: 9.383s, episode steps: 1739, steps per second: 185, episode reward: -19.000, mean reward: -0.011 [-1.000, 1.000], mean actic
1041162/1500000: episode:  679, duration: 11.455s, episode steps: 2107, steps per second: 184, episode reward: -14.000, mean reward: -0.007 [-1.000, 1.000], mean act
1043495/1500000: episode:  680, duration: 12.583s, episode steps: 2333, steps per second: 185, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean act
1045845/1500000: episode:  681, duration: 12.654s, episode steps: 2350, steps per second: 186, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean act
1047317/1500000: episode:  682, duration: 7.961s, episode steps: 1472, steps per second: 185, episode reward: -19.000, mean reward: -0.013 [-1.000, 1.000], mean acti
1049389/1500000: episode:  683, duration: 11.280s, episode steps: 2072, steps per second: 184, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean act
1051911/1500000: episode:  684, duration: 13.795s, episode steps: 2522, steps per second: 183, episode reward: -16.000, mean reward: -0.006 [-1.000, 1.000], mean act
1053866/1500000: episode:  685, duration: 10.504s, episode steps: 1955, steps per second: 186, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean act
1056280/1500000: episode:  686, duration: 13.104s, episode steps: 2414, steps per second: 184, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean act
1059551/1500000: episode:  687, duration: 17.723s, episode steps: 3271, steps per second: 185, episode reward:  -7.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1062019/1500000: episode:  688, duration: 13.440s, episode steps: 2468, steps per second: 184, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean act
1064040/1500000: episode:  689, duration: 11.042s, episode steps: 2021, steps per second: 183, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean act
1066899/1500000: episode:  690, duration: 15.540s, episode steps: 2859, steps per second: 184, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean act
1069012/1500000: episode:  691, duration: 11.527s, episode steps: 2113, steps per second: 183, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean act
1071046/1500000: episode:  692, duration: 11.001s, episode steps: 2034, steps per second: 185, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean act
1073400/1500000: episode:  693, duration: 13.023s, episode steps: 2354, steps per second: 181, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean act
1075629/1500000: episode:  694, duration: 12.144s, episode steps: 2229, steps per second: 184, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean act
1078075/1500000: episode:  695, duration: 13.737s, episode steps: 2446, steps per second: 178, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean act
1080367/1500000: episode:  696, duration: 12.632s, episode steps: 2292, steps per second: 181, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean act
1082972/1500000: episode:  697, duration: 14.132s, episode steps: 2605, steps per second: 184, episode reward: -16.000, mean reward: -0.006 [-1.000, 1.000], mean act
1084928/1500000: episode:  698, duration: 10.761s, episode steps: 1956, steps per second: 182, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean act
1087584/1500000: episode:  699, duration: 14.925s, episode steps: 2656, steps per second: 178, episode reward: -17.000, mean reward: -0.006 [-1.000, 1.000], mean act
1091193/1500000: episode:  700, duration: 19.797s, episode steps: 3609, steps per second: 182, episode reward:  -8.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1093363/1500000: episode:  701, duration: 11.837s, episode steps: 2170, steps per second: 183, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean act
1095717/1500000: episode:  702, duration: 12.989s, episode steps: 2354, steps per second: 181, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean act
1098558/1500000: episode:  703, duration: 15.296s, episode steps: 2841, steps per second: 186, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean act
1100152/1500000: episode:  704, duration: 8.610s, episode steps: 1594, steps per second: 185, episode reward: -20.000, mean reward: -0.013 [-1.000, 1.000], mean actic
1102463/1500000: episode:  705, duration: 12.553s, episode steps: 2311, steps per second: 184, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean act
1104898/1500000: episode:  706, duration: 13.227s, episode steps: 2435, steps per second: 184, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean act
1108279/1500000: episode:  707, duration: 18.564s, episode steps: 3381, steps per second: 182, episode reward:  -9.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1110656/1500000: episode:  708, duration: 12.948s, episode steps: 2377, steps per second: 184, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean act
1114117/1500000: episode:  709, duration: 18.780s, episode steps: 3461, steps per second: 184, episode reward:  -9.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1116777/1500000: episode:  710, duration: 14.738s, episode steps: 2660, steps per second: 180, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean act
1118772/1500000: episode:  711, duration: 10.877s, episode steps: 1995, steps per second: 183, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean act
1121939/1500000: episode:  712, duration: 17.213s, episode steps: 3167, steps per second: 184, episode reward:  -7.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1124336/1500000: episode:  713, duration: 13.191s, episode steps: 2397, steps per second: 182, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean act
1127423/1500000: episode:  714, duration: 16.910s, episode steps: 3087, steps per second: 183, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean act
1129440/1500000: episode:  715, duration: 11.173s, episode steps: 2017, steps per second: 181, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean act
1132326/1500000: episode:  716, duration: 15.741s, episode steps: 2886, steps per second: 183, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean act
1134701/1500000: episode:  717, duration: 12.940s, episode steps: 2375, steps per second: 184, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean act
1137943/1500000: episode:  718, duration: 17.762s, episode steps: 3242, steps per second: 183, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean act
1140884/1500000: episode:  719, duration: 16.109s, episode steps: 2941, steps per second: 183, episode reward:  -8.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1143747/1500000: episode:  720, duration: 15.988s, episode steps: 2863, steps per second: 179, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean act
1147217/1500000: episode:  721, duration: 19.048s, episode steps: 3470, steps per second: 182, episode reward:  -3.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1150382/1500000: episode:  722, duration: 17.337s, episode steps: 3165, steps per second: 183, episode reward:   4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1152700/1500000: episode:  723, duration: 12.730s, episode steps: 2318, steps per second: 182, episode reward: -13.000, mean reward: -0.006 [-1.000, 1.000], mean act
1154395/1500000: episode:  724, duration: 9.280s, episode steps: 1695, steps per second: 183, episode reward: -16.000, mean reward: -0.009 [-1.000, 1.000], mean actic
1157377/1500000: episode:  725, duration: 16.155s, episode steps: 2982, steps per second: 185, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean act
1159844/1500000: episode:  726, duration: 13.488s, episode steps: 2467, steps per second: 183, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean act
1162029/1500000: episode:  727, duration: 11.756s, episode steps: 2185, steps per second: 186, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean act
1164950/1500000: episode:  728, duration: 15.832s, episode steps: 2921, steps per second: 184, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean act
1167501/1500000: episode:  729, duration: 13.994s, episode steps: 2551, steps per second: 182, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean act
1170065/1500000: episode:  730, duration: 13.963s, episode steps: 2564, steps per second: 184, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean act
1172543/1500000: episode:  731, duration: 13.470s, episode steps: 2478, steps per second: 184, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean act
1175088/1500000: episode:  732, duration: 13.793s, episode steps: 2545, steps per second: 185, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean act
1178131/1500000: episode:  733, duration: 16.463s, episode steps: 3043, steps per second: 185, episode reward:  -9.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1180795/1500000: episode:  734, duration: 14.587s, episode steps: 2664, steps per second: 183, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean act
1183166/1500000: episode:  735, duration: 13.004s, episode steps: 2371, steps per second: 182, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean act
1186200/1500000: episode:  736, duration: 16.552s, episode steps: 3034, steps per second: 183, episode reward:  -9.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1188981/1500000: episode:  737, duration: 15.055s, episode steps: 2781, steps per second: 185, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean act
1192332/1500000: episode:  738, duration: 18.141s, episode steps: 3351, steps per second: 185, episode reward:  -6.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1195390/1500000: episode:  739, duration: 16.600s, episode steps: 3058, steps per second: 184, episode reward:  -5.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1197474/1500000: episode:  740, duration: 11.421s, episode steps: 2084, steps per second: 182, episode reward: -14.000, mean reward: -0.007 [-1.000, 1.000], mean act
1199778/1500000: episode:  741, duration: 12.815s, episode steps: 2304, steps per second: 180, episode reward: -13.000, mean reward: -0.006 [-1.000, 1.000], mean act
1202846/1500000: episode:  742, duration: 17.355s, episode steps: 3068, steps per second: 177, episode reward:  -6.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1206127/1500000: episode:  743, duration: 18.345s, episode steps: 3281, steps per second: 179, episode reward:   4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1208898/1500000: episode:  744, duration: 15.254s, episode steps: 2771, steps per second: 182, episode reward:  -8.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1210776/1500000: episode:  745, duration: 10.449s, episode steps: 1878, steps per second: 180, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean act
1213489/1500000: episode:  746, duration: 14.980s, episode steps: 2713, steps per second: 181, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean act
1216836/1500000: episode:  747, duration: 18.242s, episode steps: 3347, steps per second: 183, episode reward:  -9.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1220080/1500000: episode:  748, duration: 17.750s, episode steps: 3244, steps per second: 183, episode reward:  -9.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1223019/1500000: episode:  749, duration: 16.119s, episode steps: 2939, steps per second: 182, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean act
1225350/1500000: episode:  750, duration: 12.896s, episode steps: 2331, steps per second: 181, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean act
1228167/1500000: episode:  751, duration: 15.339s, episode steps: 2817, steps per second: 184, episode reward:  -8.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1230831/1500000: episode:  752, duration: 14.449s, episode steps: 2664, steps per second: 184, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean act
1232867/1500000: episode:  753, duration: 11.095s, episode steps: 2036, steps per second: 184, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean act
1235865/1500000: episode:  754, duration: 16.276s, episode steps: 2998, steps per second: 184, episode reward:  -9.000, mean reward: -0.003 [-1.000, 1.000], mean actic
1238875/1500000: episode:  755, duration: 16.278s, episode steps: 3010, steps per second: 185, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean act
1242602/1500000: episode:  756, duration: 20.379s, episode steps: 3727, steps per second: 183, episode reward:  -2.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1245707/1500000: episode:  757, duration: 16.780s, episode steps: 3105, steps per second: 185, episode reward:  -6.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1249655/1500000: episode:  758, duration: 21.428s, episode steps: 3948, steps per second: 184, episode reward:  -3.000, mean reward: -0.001 [-1.000, 1.000], mean actic
1252801/1500000: episode:  759, duration: 17.018s, episode steps: 3146, steps per second: 185, episode reward:   3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1255835/1500000: episode:  760, duration: 16.656s, episode steps: 3034, steps per second: 182, episode reward:  -7.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1258877/1500000: episode:  761, duration: 16.749s, episode steps: 3042, steps per second: 182, episode reward:  -5.000, mean reward: -0.002 [-1.000, 1.000], mean actic
1261482/1500000: episode:  762, duration: 14.064s, episode steps: 2605, steps per second: 185, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
```
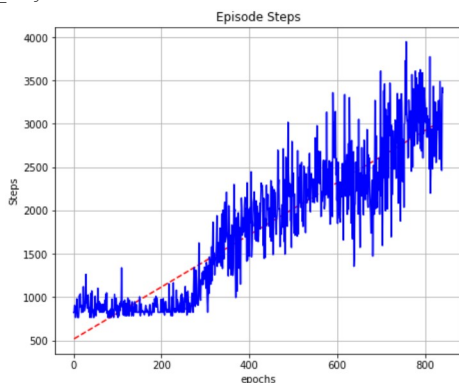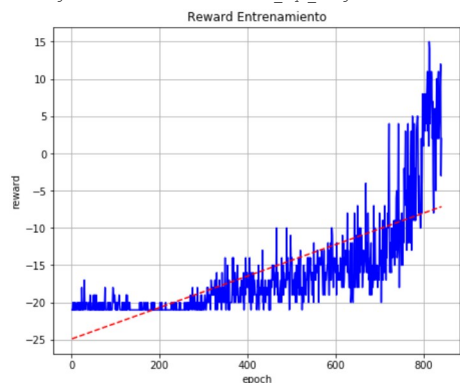
```
1264106/1500000: episode: 763, duration: 14.181s, episode steps: 2624, steps per second: 185, episode reward: -7.000, mean reward: -0.003 [-1.000, 1.000], mean actio
1267379/1500000: episode: 764, duration: 17.713s, episode steps: 3273, steps per second: 185, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean actio
1270504/1500000: episode: 765, duration: 17.164s, episode steps: 3125, steps per second: 182, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1273343/1500000: episode: 766, duration: 15.476s, episode steps: 2839, steps per second: 183, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1276408/1500000: episode: 767, duration: 16.664s, episode steps: 3065, steps per second: 184, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean actio
1279547/1500000: episode: 768, duration: 16.951s, episode steps: 3139, steps per second: 185, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean actio
1282696/1500000: episode: 769, duration: 17.073s, episode steps: 3149, steps per second: 184, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean actio
1285148/1500000: episode: 770, duration: 13.449s, episode steps: 2452, steps per second: 182, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
1288717/1500000: episode: 771, duration: 19.242s, episode steps: 3569, steps per second: 185, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean actio
1291798/1500000: episode: 772, duration: 16.675s, episode steps: 3081, steps per second: 185, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1294300/1500000: episode: 773, duration: 13.533s, episode steps: 2502, steps per second: 185, episode reward: -13.000, mean reward: -0.005 [-1.000, 1.000], mean acti
1297705/1500000: episode: 774, duration: 18.601s, episode steps: 3405, steps per second: 183, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean actio
1301048/1500000: episode: 775, duration: 18.116s, episode steps: 3343, steps per second: 185, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1303794/1500000: episode: 776, duration: 14.836s, episode steps: 2746, steps per second: 185, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean actio
1306750/1500000: episode: 777, duration: 16.079s, episode steps: 2956, steps per second: 184, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean actio
1310037/1500000: episode: 778, duration: 17.786s, episode steps: 3287, steps per second: 185, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean actio
1313646/1500000: episode: 779, duration: 19.878s, episode steps: 3609, steps per second: 182, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1316561/1500000: episode: 780, duration: 15.814s, episode steps: 2915, steps per second: 184, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean actio
1320093/1500000: episode: 781, duration: 18.984s, episode steps: 3532, steps per second: 186, episode reward: -2.000, mean reward: -0.001 [-1.000, 1.000], mean actio
1323124/1500000: episode: 782, duration: 16.147s, episode steps: 3031, steps per second: 188, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean actio
1325890/1500000: episode: 783, duration: 14.899s, episode steps: 2766, steps per second: 186, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean actio
1329342/1500000: episode: 784, duration: 18.580s, episode steps: 3452, steps per second: 186, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1332775/1500000: episode: 785, duration: 18.360s, episode steps: 3433, steps per second: 187, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1335848/1500000: episode: 786, duration: 16.482s, episode steps: 3073, steps per second: 186, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1338864/1500000: episode: 787, duration: 16.216s, episode steps: 3016, steps per second: 186, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1342454/1500000: episode: 788, duration: 19.334s, episode steps: 3590, steps per second: 186, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean actio
1345253/1500000: episode: 789, duration: 15.319s, episode steps: 2799, steps per second: 183, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean actio
1348877/1500000: episode: 790, duration: 19.530s, episode steps: 3624, steps per second: 186, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean actio
1351589/1500000: episode: 791, duration: 14.597s, episode steps: 2712, steps per second: 186, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean actio
1354484/1500000: episode: 792, duration: 15.665s, episode steps: 2895, steps per second: 185, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean actio
1357170/1500000: episode: 793, duration: 14.518s, episode steps: 2686, steps per second: 185, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1359954/1500000: episode: 794, duration: 15.260s, episode steps: 2784, steps per second: 182, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
1363543/1500000: episode: 795, duration: 19.335s, episode steps: 3589, steps per second: 186, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1366148/1500000: episode: 796, duration: 13.995s, episode steps: 2605, steps per second: 186, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean actio
1369295/1500000: episode: 797, duration: 16.990s, episode steps: 3147, steps per second: 185, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean actio
1372183/1500000: episode: 798, duration: 16.093s, episode steps: 2888, steps per second: 179, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1375599/1500000: episode: 799, duration: 18.680s, episode steps: 3416, steps per second: 183, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1378904/1500000: episode: 800, duration: 17.802s, episode steps: 3305, steps per second: 186, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
1382129/1500000: episode: 801, duration: 17.531s, episode steps: 3225, steps per second: 184, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1385144/1500000: episode: 802, duration: 16.284s, episode steps: 3015, steps per second: 185, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1388282/1500000: episode: 803, duration: 17.043s, episode steps: 3138, steps per second: 184, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1391466/1500000: episode: 804, duration: 17.308s, episode steps: 3184, steps per second: 184, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1394313/1500000: episode: 805, duration: 15.309s, episode steps: 2847, steps per second: 186, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1397407/1500000: episode: 806, duration: 16.731s, episode steps: 3094, steps per second: 185, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1400471/1500000: episode: 807, duration: 16.921s, episode steps: 3064, steps per second: 181, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1403351/1500000: episode: 808, duration: 15.898s, episode steps: 2880, steps per second: 181, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1406192/1500000: episode: 809, duration: 15.449s, episode steps: 2841, steps per second: 184, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean actio
1409247/1500000: episode: 810, duration: 16.814s, episode steps: 3055, steps per second: 182, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1411726/1500000: episode: 811, duration: 13.483s, episode steps: 2479, steps per second: 184, episode reward: 9.000, mean reward: 0.004 [-1.000, 1.000], mean action:
1415501/1500000: episode: 812, duration: 20.474s, episode steps: 3775, steps per second: 184, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
1417700/1500000: episode: 813, duration: 11.990s, episode steps: 2199, steps per second: 183, episode reward: 15.000, mean reward: 0.007 [-1.000, 1.000], mean actio
1420099/1500000: episode: 814, duration: 12.959s, episode steps: 2399, steps per second: 185, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean actio
1423029/1500000: episode: 815, duration: 15.760s, episode steps: 2930, steps per second: 186, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1425968/1500000: episode: 816, duration: 15.813s, episode steps: 2939, steps per second: 186, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1428791/1500000: episode: 817, duration: 15.589s, episode steps: 2823, steps per second: 181, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1431272/1500000: episode: 818, duration: 13.405s, episode steps: 2481, steps per second: 185, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean actio
1434050/1500000: episode: 819, duration: 15.265s, episode steps: 2778, steps per second: 182, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean actio
1437484/1500000: episode: 820, duration: 18.531s, episode steps: 3434, steps per second: 185, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1440136/1500000: episode: 821, duration: 14.429s, episode steps: 2652, steps per second: 184, episode reward: 7.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1442758/1500000: episode: 822, duration: 14.418s, episode steps: 2622, steps per second: 182, episode reward: -4.000, mean reward: -0.002 [-1.000, 1.000], mean actio
1445411/1500000: episode: 823, duration: 14.666s, episode steps: 2653, steps per second: 181, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean actio
1448662/1500000: episode: 824, duration: 17.731s, episode steps: 3251, steps per second: 183, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1451419/1500000: episode: 825, duration: 14.842s, episode steps: 2757, steps per second: 186, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1453972/1500000: episode: 826, duration: 13.922s, episode steps: 2553, steps per second: 183, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1456652/1500000: episode: 827, duration: 14.448s, episode steps: 2680, steps per second: 185, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1459593/1500000: episode: 828, duration: 15.805s, episode steps: 2941, steps per second: 186, episode reward: -5.000, mean reward: -0.002 [-1.000, 1.000], mean actio
1462733/1500000: episode: 829, duration: 17.168s, episode steps: 3140, steps per second: 183, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1465667/1500000: episode: 830, duration: 16.005s, episode steps: 2934, steps per second: 183, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actio
1468663/1500000: episode: 831, duration: 16.487s, episode steps: 2996, steps per second: 182, episode reward: 8.000, mean reward: 0.003 [-1.000, 1.000], mean action:
1471933/1500000: episode: 832, duration: 17.705s, episode steps: 3270, steps per second: 185, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1474618/1500000: episode: 833, duration: 14.475s, episode steps: 2685, steps per second: 185, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean actio
1477208/1500000: episode: 834, duration: 14.136s, episode steps: 2590, steps per second: 183, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean actio
1480697/1500000: episode: 835, duration: 18.887s, episode steps: 3489, steps per second: 185, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1483897/1500000: episode: 836, duration: 17.426s, episode steps: 3200, steps per second: 184, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
1486841/1500000: episode: 837, duration: 16.031s, episode steps: 2944, steps per second: 184, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
1489414/1500000: episode: 838, duration: 13.853s, episode steps: 2573, steps per second: 186, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean actio
1491875/1500000: episode: 839, duration: 13.318s, episode steps: 2461, steps per second: 185, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean actio
1495294/1500000: episode: 840, duration: 18.356s, episode steps: 3419, steps per second: 186, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean actio
1498659/1500000: episode: 841, duration: 18.020s, episode steps: 3365, steps per second: 187, episode reward: 2.000, mean reward: 0.001 [-1.000, 1.000], mean action:
done, took 8225.261 seconds
Pesos guardados en: 202001100948_dqn_PongDeterministic-v0_weights.h5f
```


Reward Entrenamiento / Episode Steps

```
Testing for 10 episodes ...
Episode 1: reward: 14.000, steps: 2287
Episode 2: reward: 8.000, steps: 2776
Episode 3: reward: 14.000, steps: 2335
Episode 4: reward: 12.000, steps: 2390
Episode 5: reward: 14.000, steps: 2332
Episode 6: reward: 9.000, steps: 2696
```

```
Episode 7: reward: 1.000, steps: 3392
Episode 8: reward: 6.000, steps: 3212
Episode 9: reward: 7.000, steps: 2672
Episode 10: reward: -3.000, steps: 3007
mean reward: 8.2
mean_steps: 2709.9
(8.2, 2709.9)
```