

Training for 5000000 steps ...

| | | | | | | | | |
|----------------|--------------|-------------------|----------------------|------------------------|--------------------------|---------------------|------------------|--------------|
| 1646/5000000: | episode: 2, | duration: 2.34, | episode steps: 824, | steps per second: 352, | episode reward: -21.000, | mean reward: -0.025 | [-1.000, 0.000], | mean action: |
| 2526/5000000: | episode: 3, | duration: 2.452s, | episode steps: 880, | steps per second: 359, | episode reward: -21.000, | mean reward: -0.024 | [-1.000, 0.000], | mean action: |
| 3447/5000000: | episode: 4, | duration: 2.642s, | episode steps: 921, | steps per second: 349, | episode reward: -20.000, | mean reward: -0.022 | [-1.000, 1.000], | mean action: |
| 4271/5000000: | episode: 5, | duration: 2.413s, | episode steps: 824, | steps per second: 341, | episode reward: -21.000, | mean reward: -0.025 | [-1.000, 0.000], | mean action: |
| 5035/5000000: | episode: 6, | duration: 2.184s, | episode steps: 764, | steps per second: 350, | episode reward: -21.000, | mean reward: -0.027 | [-1.000, 0.000], | mean action: |
| 5859/5000000: | episode: 7, | duration: 2.348s, | episode steps: 824, | steps per second: 351, | episode reward: -21.000, | mean reward: -0.025 | [-1.000, 0.000], | mean action: |
| 6870/5000000: | episode: 8, | duration: 2.911s, | episode steps: 1011, | steps per second: 347, | episode reward: -20.000, | mean reward: -0.020 | [-1.000, 1.000], | mean action: |
| 7670/5000000: | episode: 9, | duration: 2.276s, | episode steps: 800, | steps per second: 352, | episode reward: -21.000, | mean reward: -0.026 | [-1.000, 0.000], | mean action: |
| 8462/5000000: | episode: 10, | duration: 2.317s, | episode steps: 792, | steps per second: 342, | episode reward: -21.000, | mean reward: -0.027 | [-1.000, 0.000], | mean action: |
| 9347/5000000: | episode: 11, | duration: 2.570s, | episode steps: 885, | steps per second: 344, | episode reward: -21.000, | mean reward: -0.024 | [-1.000, 0.000], | mean action: |
| 10171/5000000: | episode: 12, | duration: 4.930s, | episode steps: 824, | steps per second: 167, | episode reward: -21.000, | mean reward: -0.025 | [-1.000, 0.000], | mean action: |
| 11041/5000000: | episode: 13, | duration: 5.799s, | episode steps: 870, | steps per second: 150, | episode reward: -20.000, | mean reward: -0.023 | [-1.000, 1.000], | mean action: |
| 11921/5000000: | episode: 14, | duration: 5.826s, | episode steps: 880, | steps per second: 151, | episode reward: -21.000, | mean reward: -0.024 | [-1.000, 0.000], | mean action: |
| 12915/5000000: | episode: 15, | duration: 6.507s, | episode steps: 994, | steps per second: 153, | episode reward: -20.000, | mean reward: -0.020 | [-1.000, 1.000], | mean action: |
| 14013/5000000: | episode: 16, | duration: 7.360s, | episode steps: 1098, | steps per second: 149, | episode reward: -20.000, | mean reward: -0.018 | [-1.000, 1.000], | mean action: |
| 14856/5000000: | episode: 17, | duration: 5.710s, | episode steps: 843, | steps per second: 148, | episode reward: -21.000, | mean reward: -0.025 | [-1.000, 0.000], | mean action: |
| 15802/5000000: | episode: 18, | duration: 6.529s, | episode steps: 946, | steps per second: 145, | episode reward: -21.000, | mean reward: -0.022 | [-1.000, 0.000], | mean action: |
| 16672/5000000: | episode: 19, | duration: 5.834s, | episode steps: 870, | steps per second: 149, | episode reward: -20.000, | mean reward: -0.023 | [-1.000, 1.000], | mean action: |
| 17542/5000000: | episode: 20, | duration: 5.857s, | episode steps: 870, | steps per second: 149, | episode reward: -20.000, | mean reward: -0.023 | [-1.000, 1.000], | mean action: |
| 18473/5000000: | episode: 21, | duration: 6.149s, | episode steps: 931, | steps per second: 151, | episode reward: -20.000, | mean reward: -0.021 | [-1.000, 1.000], | mean action: |
| 19297/5000000: | episode: 22, | duration: 5.358s, | episode steps: 824, | steps per second: 154, | episode reward: -21.000, | mean reward: -0.025 | [-1.000, 0.000], | mean action: |
| 20308/5000000: | episode: 23, | duration: 6.711s, | episode steps: 1011, | steps per second: 151, | episode reward: -21.000, | mean reward: -0.021 | [-1.000, 0.000], | mean action: |
| 21348/5000000: | episode: 24, | duration: 6.890s, | episode steps: 1040, | steps per second: 151, | episode reward: -20.000, | mean reward: -0.019 | [-1.000, 1.000], | mean action: |
| 22202/5000000: | episode: 25, | duration: 5.643s, | episode steps: 854, | steps per second: 151, | episode reward: -21.000, | mean reward: -0.025 | [-1.000, 0.000], | mean action: |
| 23132/5000000: | episode: 26, | duration: 6.146s, | episode steps: 930, | steps per second: 151, | episode reward: -20.000, | mean reward: -0.022 | [-1.000, 1.000], | mean action: |
| 24147/5000000: | episode: 27, | duration: 6.668s, | episode steps: 1015, | steps per second: 152, | episode reward: -20.000, | mean reward: -0.020 | [-1.000, 1.000], | mean action: |
| 25017/5000000: | episode: 28, | duration: 5.781s, | episode steps: 870, | steps per second: 151, | episode reward: -20.000, | mean reward: -0.023 | [-1.000, 1.000], | mean action: |
| 25781/5000000: | episode: 29, | duration: 5.171s, | episode steps: 764, | steps per second: 148, | episode reward: -21.000, | mean reward: -0.027 | [-1.000, 0.000], | mean action: |
| 26748/5000000: | episode: 30, | duration: 6.430s, | episode steps: 967, | steps per second: 150, | episode reward: -20.000, | mean reward: -0.021 | [-1.000, 1.000], | mean action: |
| 27696/5000000: | episode: 31, | duration: 6.166s, | episode steps: 948, | steps per second: 154, | episode reward: -20.000, | mean reward: -0.021 | [-1.000, 1.000], | mean action: |
| 28539/5000000: | episode: 32, | duration: 5.456s, | episode steps: 843, | steps per second: 154, | episode reward: -20.000, | mean reward: -0.024 | [-1.000, 1.000], | mean action: |
| 29451/5000000: | episode: 33, | duration: 6.057s, | episode steps: 912, | steps per second: 151, | episode reward: -21.000, | mean reward: -0.023 | [-1.000, 0.000], | mean action: |
| 30383/5000000: | episode: 34, | duration: 6.064s, | episode steps: 932, | steps per second: 154, | episode reward: -20 | | | |

105382/5000000: episode: 109, duration: 12.852s, episode steps: 1997, steps per second: 155, episode reward: -14.000, mean reward: -0.007 [-1.000, 1.000], mean acti
107023/5000000: episode: 110, duration: 10.483s, episode steps: 1641, steps per second: 157, episode reward: -16.000, mean reward: -0.010 [-1.000, 1.000], mean acti
108429/5000000: episode: 111, duration: 9.088s, episode steps: 1406, steps per second: 155, episode reward: -19.000, mean reward: -0.014 [-1.000, 1.000], mean acti
109421/5000000: episode: 112, duration: 6.432s, episode steps: 992, steps per second: 154, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean acti
110830/5000000: episode: 113, duration: 9.564s, episode steps: 1409, steps per second: 147, episode reward: -17.000, mean reward: -0.012 [-1.000, 1.000], mean acti
112462/5000000: episode: 114, duration: 10.550s, episode steps: 1632, steps per second: 155, episode reward: -16.000, mean reward: -0.010 [-1.000, 1.000], mean acti
113857/5000000: episode: 115, duration: 9.007s, episode steps: 1395, steps per second: 155, episode reward: -18.000, mean reward: -0.013 [-1.000, 1.000], mean acti
115176/5000000: episode: 116, duration: 8.578s, episode steps: 1319, steps per second: 154, episode reward: -19.000, mean reward: -0.014 [-1.000, 1.000], mean acti
116747/5000000: episode: 117, duration: 10.256s, episode steps: 1571, steps per second: 153, episode reward: -18.000, mean reward: -0.011 [-1.000, 1.000], mean acti
118601/5000000: episode: 118, duration: 12.070s, episode steps: 1854, steps per second: 154, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean acti
120709/5000000: episode: 119, duration: 13.802s, episode steps: 2108, steps per second: 153, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
122358/5000000: episode: 120, duration: 10.687s, episode steps: 1649, steps per second: 154, episode reward: -19.000, mean reward: -0.012 [-1.000, 1.000], mean acti
124489/5000000: episode: 121, duration: 13.833s, episode steps: 2131, steps per second: 154, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
126670/5000000: episode: 122, duration: 14.061s, episode steps: 2181, steps per second: 155, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
128105/5000000: episode: 123, duration: 9.358s, episode steps: 1435, steps per second: 153, episode reward: -18.000, mean reward: -0.013 [-1.000, 1.000], mean acti
130053/5000000: episode: 124, duration: 12.590s, episode steps: 1948, steps per second: 155, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
132351/5000000: episode: 125, duration: 14.847s, episode steps: 2298, steps per second: 155, episode reward: -13.000, mean reward: -0.006 [-1.000, 1.000], mean acti
134775/5000000: episode: 126, duration: 15.541s, episode steps: 2424, steps per second: 156, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
136526/5000000: episode: 127, duration: 11.297s, episode steps: 1751, steps per second: 155, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean acti
13955/5000000: episode: 128, duration: 9.213s, episode steps: 1429, steps per second: 155, episode reward: -20.000, mean reward: -0.014 [-1.000, 1.000], mean acti
139704/5000000: episode: 129, duration: 11.263s, episode steps: 1749, steps per second: 155, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean acti
141733/5000000: episode: 130, duration: 13.012s, episode steps: 2029, steps per second: 156, episode reward: -19.000, mean reward: -0.009 [-1.000, 1.000], mean acti
143680/5000000: episode: 131, duration: 12.569s, episode steps: 1947, steps per second: 155, episode reward: -16.000, mean reward: -0.008 [-1.000, 1.000], mean acti
146004/5000000: episode: 132, duration: 15.039s, episode steps: 2324, steps per second: 155, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean acti
147483/5000000: episode: 133, duration: 9.553s, episode steps: 1479, steps per second: 155, episode reward: -21.000, mean reward: -0.014 [-1.000, 0.000], mean acti
149445/5000000: episode: 134, duration: 12.724s, episode steps: 1962, steps per second: 154, episode reward: -19.000, mean reward: -0.010 [-1.000, 1.000], mean acti
150385/5000000: episode: 135, duration: 6.075s, episode steps: 940, steps per second: 155, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean acti
152150/5000000: episode: 136, duration: 11.310s, episode steps: 1765, steps per second: 156, episode reward: -20.000, mean reward: -0.011 [-1.000, 1.000], mean acti
154456/5000000: episode: 137, duration: 14.734s, episode steps: 2306, steps per second: 157, episode reward: -20.000, mean reward: -0.009 [-1.000, 1.000], mean acti
156693/5000000: episode: 138, duration: 14.439s, episode steps: 2237, steps per second: 155, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean acti
159282/5000000: episode: 139, duration: 17.055s, episode steps: 2589, steps per second: 152, episode reward: -19.000, mean reward: -0.007 [-1.000, 1.000], mean acti
161704/5000000: episode: 140, duration: 15.585s, episode steps: 2422, steps per second: 155, episode reward: -19.000, mean reward: -0.008 [-1.000, 1.000], mean acti
163799/5000000: episode: 141, duration: 13.461s, episode steps: 2095, steps per second: 156, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean acti
165783/5000000: episode: 142, duration: 12.828s, episode steps: 1984, steps per second: 155, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
167580/5000000: episode: 143, duration: 11.663s, episode steps: 1797, steps per second: 154, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean acti
169582/5000000: episode: 144, duration: 12.754s, episode steps: 2002, steps per second: 157, episode reward: -20.000, mean reward: -0.010 [-1.000, 1.000], mean acti
172240/5000000: episode: 145, duration: 17.083s, episode steps: 2658, steps per second: 156, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
175029/5000000: episode: 146, duration: 17.859s, episode steps: 2789, steps per second: 156, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean acti
177177/5000000: episode: 147, duration: 13.740s, episode steps: 2148, steps per second: 156, episode reward: -19.000, mean reward: -0.009 [-1.000, 1.000], mean acti
179827/5000000: episode: 148, duration: 16.859s, episode steps: 2650, steps per second: 157, episode reward: -16.000, mean reward: -0.006 [-1.000, 1.000], mean acti
182202/5000000: episode: 149, duration: 15.184s, episode steps: 2375, steps per second: 156, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean acti
184536/5000000: episode: 150, duration: 15.045s, episode steps: 2334, steps per second: 155, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
186410/5000000: episode: 151, duration: 12.098s, episode steps: 1874, steps per second: 155, episode reward: -16.000, mean reward: -0.009 [-1.000, 1.000], mean acti
188987/5000000: episode: 152, duration: 16.685s, episode steps: 2577, steps per second: 154, episode reward: -18.000, mean reward: -0.007 [-1.000, 1.000], mean acti
191427/5000000: episode: 153, duration: 15.790s, episode steps: 2440, steps per second: 155, episode reward: -18.000, mean reward: -0.007 [-1.000, 1.000], mean acti
193531/5000000: episode: 154, duration: 13.562s, episode steps: 2104, steps per second: 155, episode reward: -19.000, mean reward: -0.009 [-1.000, 1.000], mean acti
196149/5000000: episode: 155, duration: 17.036s, episode steps: 2618, steps per second: 154, episode reward: -20.000, mean reward: -0.008 [-1.000, 1.000], mean acti
199277/5000000: episode: 156, duration: 20.204s, episode steps: 3128, steps per second: 155, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
200740/5000000: episode: 157, duration: 9.488s, episode steps: 1463, steps per second: 154, episode reward: -20.000, mean reward: -0.014 [-1.000, 1.000], mean acti
204096/5000000: episode: 158, duration: 21.829s, episode steps: 3356, steps per second: 154, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
205907/5000000: episode: 159, duration: 11.887s, episode steps: 1811, steps per second: 152, episode reward: -16.000, mean reward: -0.009 [-1.000, 1.000], mean acti
208291/5000000: episode: 160, duration: 15.500s, episode steps: 2384, steps per second: 154, episode reward: -18.000, mean reward: -0.008 [-1.000, 1.000], mean acti
210481/5000000: episode: 161, duration: 14.196s, episode steps: 2190, steps per second: 154, episode reward: -20.000, mean reward: -0.009 [-1.000, 1.000], mean acti
212534/5000000: episode: 162, duration: 13.394s, episode steps: 2053, steps per second: 153, episode reward: -21.000, mean reward: -0.010 [-1.000, 0.000], mean acti
214019/5000000: episode: 163, duration: 9.803s, episode steps: 1485, steps per second: 151, episode reward: -21.000, mean reward: -0.014 [-1.000, 0.000], mean acti
217166/5000000: episode: 164, duration: 20.369s, episode steps: 3147, steps per second: 154, episode reward: -14.000, mean reward: -0.004 [-1.000, 1.000], mean acti
219885/5000000: episode: 165, duration: 17.442s, episode steps: 2719, steps per second: 156, episode reward: -17.000, mean reward: -0.006 [-1.000, 1.000], mean acti
223356/5000000: episode: 166, duration: 22.132s, episode steps: 3471, steps per second: 157, episode reward: -14.000, mean reward: -0.004 [-1.000, 1.000], mean acti
227430/5000000: episode: 167, duration: 25.999s, episode steps: 4074, steps per second: 157, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
229605/5000000: episode: 168, duration: 13.885s, episode steps: 2175, steps per second: 157, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean acti
232511/5000000: episode: 169, duration: 18.588s, episode steps: 2906, steps per second: 156, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
235236/5000000: episode: 170, duration: 17.429s, episode steps: 2725, steps per second: 156, episode reward: -20.000, mean reward: -0.007 [-1.000, 1.000], mean acti
238307/5000000: episode: 171, duration: 19.647s, episode steps: 3071, steps per second: 156, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean acti
241195/5000000: episode: 172, duration: 18.517s, episode steps: 2888, steps per second: 156, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean acti
244191/5000000: episode: 173, duration: 19.118s, episode steps: 2996, steps per second: 157, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean acti
247290/5000000: episode: 174, duration: 19.676s, episode steps: 3099, steps per second: 158, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean acti
250052/5000000: episode: 175, duration: 17.761s, episode steps: 2762, steps per second: 156, episode reward: -18.000, mean reward: -0.007 [-1.000, 1.000], mean acti
253277/5000000: episode: 176, duration: 20.584s, episode steps: 3225, steps per second: 157, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
257117/5000000: episode: 177, duration: 24.914s, episode steps: 3840, steps per second: 154, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean acti
260403/5000000: episode: 178, duration: 21.165s, episode steps: 3286, steps per second: 155, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
263172/5000000: episode: 179, duration: 17.859s, episode steps: 2769, steps per second: 155, episode reward: 17.000, mean reward: 0.006 [-1.000, 1.000], mean acti
266635/5000000: episode: 180, duration: 22.128s, episode steps: 3463, steps per second: 156, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
270450/5000000: episode: 181, duration: 24.363s, episode steps: 3815, steps per second: 157, episode reward: -9.000, mean reward: -0.002 [-1.000, 1.000], mean acti
273287/5000000: episode: 182, duration: 18.074s, episode steps: 2837, steps per second: 157, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean acti
277321/5000000: episode: 183, duration: 25.702s, episode steps: 4034, steps per second: 157, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean acti
279974/5000000: episode: 184, duration: 16.867s, episode steps: 2653, steps per second: 155, episode reward: -17.000, mean reward: -0.006 [-1.000, 1.000], mean acti
284171/5000000: episode: 185, duration: 26.630s, episode steps: 4197, steps per second: 158, episode reward: -6.000, mean reward: -0.001 [-1.000, 1.000], mean acti
287468/5000000: episode: 186, duration: 21.105s, episode steps: 3297, steps per second: 156, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
291761/5000000: episode: 187, duration: 27.373s, episode steps: 4293, steps per second: 157, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean acti
295134/5000000: episode: 188, duration: 21.619s, episode steps: 3373, steps per second: 156, episode reward: -14.000, mean reward: -0.004 [-1.000, 1.000], mean acti
297562/5000000: episode: 189, duration: 15.472s, episode steps: 2428, steps per second: 157, episode reward: -18.000, mean reward: -0.007 [-1.000, 1.000], mean acti
300494/5000000: episode: 190, duration: 18.761s, episode steps: 2932, steps per second: 156, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
303160/5000000: episode: 191, duration: 17.508s, episode steps: 2666, steps per second: 152, episode reward: -20.000, mean reward: -0.008 [-1.000, 1.000], mean acti
307255/5000000: episode: 192, duration: 26.129s, episode steps: 4095, steps per second: 157, episode reward: -6.000, mean reward: -0.001 [-1.000, 1.000], mean acti
310720/5000000: episode: 193, duration: 22.231s, episode steps: 3465, steps per second: 156, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean acti
314110/5000000: episode: 194, duration: 21.506s, episode steps: 3390, steps per second: 158, episode reward: -15.000, mean reward: -0.004 [-1.000, 1.000], mean acti
316655/5000000: episode: 195, duration: 16.246s, episode steps: 2545, steps per second: 157, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean acti
321081/5000000: episode: 196, duration: 28.165s, episode steps: 4426, steps per second: 157, episode reward: 2.000, mean reward: 0.000 [-1.000, 1.000], mean action:
323848/5000000: episode: 197, duration: 17.682s, episode steps: 2767, steps per second: 156, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean acti
326759/5000000: episode: 198, duration: 18.398s, episode steps: 2911, steps per second: 158, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
329015/5000000: episode: 199, duration: 14.224s, episode steps: 2256, steps per second: 159, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean acti
332172/5000000: episode: 200, duration: 20.064s, episode steps: 3157, steps per second: 157, episode reward: -18.000, mean reward: -0.006 [-1.000, 1.000], mean acti
336718/5000000: episode: 201, duration: 28.817s, episode steps: 4546, steps per second: 158, episode reward: -2.000, mean reward: -0.000 [-1.000, 1.000], mean acti
339480/5000000: episode: 202, duration: 17.568s, episode steps: 2762, steps per second: 157, episode reward: -19.000, mean reward: -0.007 [-1.000, 1.000], mean acti
343678/5000000: episode: 203, duration: 26.590s, episode steps: 4198, steps per second: 158, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti
347167/5000000: episode: 204, duration: 22.220s, episode steps: 3489, steps per second: 157, episode reward: -16.000, mean reward: -0.005 [-1.000, 1.000], mean acti
350964/5000000: episode: 205, duration: 24.143s, episode steps: 3797, steps per second: 157, episode reward: -14.000, mean reward: -0.004 [-1.000, 1.000], mean acti
355155/5000000: episode: 206, duration: 26.874s, episode steps: 4191, steps per second: 156, episode reward: -8.000, mean reward: -0.002 [-1.000, 1.000], mean acti
357575/5000000: episode: 207, duration: 15.551s, episode steps: 2420, steps per second: 156, episode reward: -19.000, mean reward: -0.008 [-1.000, 1.000], mean acti
360610/5000000: episode: 208, duration: 19.179s, episode steps: 3035, steps per second: 158, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean acti
364492/5000000: episode: 209, duration: 24.456s, episode steps: 3882, steps per second: 159, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean acti
368477/5000000: episode: 210, duration: 25.057s, episode steps: 3985, steps per second: 159, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
371081/5000000: episode: 211, duration: 16.455s, episode steps: 2604, steps per second: 158, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean acti
376354/5000000: episode: 212, duration: 33.293s, episode steps: 5273, steps per second: 158, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean acti
381516/5000000: episode: 213, duration: 32.938s, episode steps: 5162, steps per second: 157, episode reward: -11.000, mean reward: -0.002 [-1.000, 1.000], mean acti
386332/5000000: episode: 214, duration: 30.843s, episode steps: 4816, steps per second: 156, episode reward: -10.000, mean reward: -0.002 [-1.000, 1.000], mean acti
391027/5000000: episode: 215, duration: 29.843s, episode steps: 4695, steps per second: 157, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
395385/5000000: episode: 216, duration: 27.673s, episode steps: 4358, steps per second: 157, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
399093/5000000: episode: 217, duration: 23.586s, episode steps: 3708, steps per second: 157, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean acti

```

402884/5000000: episode: 218, duration: 24.411s, episode steps: 3791, steps per second: 155, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
407466/5000000: episode: 219, duration: 29.316s, episode steps: 4582, steps per second: 156, episode reward: -12.000, mean reward: -0.003 [-1.000, 1.000], mean action:
410594/5000000: episode: 220, duration: 20.010s, episode steps: 3128, steps per second: 156, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean action:
414053/5000000: episode: 221, duration: 22.146s, episode steps: 3459, steps per second: 156, episode reward: -16.000, mean reward: -0.005 [-1.000, 1.000], mean action:
418620/5000000: episode: 222, duration: 29.138s, episode steps: 4567, steps per second: 157, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean action:
421589/5000000: episode: 223, duration: 18.992s, episode steps: 2969, steps per second: 156, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean action:
423984/5000000: episode: 224, duration: 15.151s, episode steps: 2395, steps per second: 158, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean action:
425587/5000000: episode: 225, duration: 10.270s, episode steps: 1603, steps per second: 156, episode reward: -21.000, mean reward: -0.013 [-1.000, 0.000], mean action:
429707/5000000: episode: 226, duration: 26.308s, episode steps: 4120, steps per second: 157, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean action:
434796/5000000: episode: 227, duration: 32.506s, episode steps: 5089, steps per second: 157, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean action:
438739/5000000: episode: 228, duration: 25.188s, episode steps: 3943, steps per second: 157, episode reward: -14.000, mean reward: -0.004 [-1.000, 1.000], mean action:
443196/5000000: episode: 229, duration: 28.533s, episode steps: 4457, steps per second: 156, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean action:
448404/5000000: episode: 230, duration: 33.750s, episode steps: 5244, steps per second: 155, episode reward: -6.000, mean reward: -0.001 [-1.000, 1.000], mean action:
452167/5000000: episode: 231, duration: 23.989s, episode steps: 3727, steps per second: 155, episode reward: -13.000, mean reward: -0.003 [-1.000, 1.000], mean action:
455473/5000000: episode: 232, duration: 21.291s, episode steps: 3306, steps per second: 155, episode reward: -16.000, mean reward: -0.005 [-1.000, 1.000], mean action:
459551/5000000: episode: 233, duration: 25.977s, episode steps: 4078, steps per second: 157, episode reward: -8.000, mean reward: -0.002 [-1.000, 1.000], mean action:
465287/5000000: episode: 234, duration: 36.642s, episode steps: 5736, steps per second: 157, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean action:
469596/5000000: episode: 235, duration: 27.613s, episode steps: 4309, steps per second: 156, episode reward: -9.000, mean reward: -0.002 [-1.000, 1.000], mean action:
473991/5000000: episode: 236, duration: 28.168s, episode steps: 4395, steps per second: 156, episode reward: -6.000, mean reward: -0.001 [-1.000, 1.000], mean action:
478404/5000000: episode: 237, duration: 28.249s, episode steps: 4413, steps per second: 156, episode reward: -10.000, mean reward: -0.002 [-1.000, 1.000], mean action:
483224/5000000: episode: 238, duration: 30.750s, episode steps: 4820, steps per second: 157, episode reward: 2.000, mean reward: 0.000 [-1.000, 1.000], mean action:
487554/5000000: episode: 239, duration: 27.642s, episode steps: 4330, steps per second: 157, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
490812/5000000: episode: 240, duration: 20.872s, episode steps: 3258, steps per second: 156, episode reward: -14.000, mean reward: -0.004 [-1.000, 1.000], mean action:
494043/5000000: episode: 241, duration: 20.766s, episode steps: 3231, steps per second: 156, episode reward: -18.000, mean reward: -0.006 [-1.000, 1.000], mean action:
498762/5000000: episode: 242, duration: 30.622s, episode steps: 4719, steps per second: 154, episode reward: -7.000, mean reward: -0.001 [-1.000, 1.000], mean action:
501695/5000000: episode: 243, duration: 19.099s, episode steps: 2933, steps per second: 154, episode reward: -16.000, mean reward: -0.005 [-1.000, 1.000], mean action:
506963/5000000: episode: 244, duration: 33.882s, episode steps: 5268, steps per second: 155, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
510666/5000000: episode: 245, duration: 23.688s, episode steps: 3703, steps per second: 156, episode reward: -15.000, mean reward: -0.004 [-1.000, 1.000], mean action:
514530/5000000: episode: 246, duration: 24.602s, episode steps: 3864, steps per second: 157, episode reward: -8.000, mean reward: -0.002 [-1.000, 1.000], mean action:
520300/5000000: episode: 247, duration: 36.839s, episode steps: 5770, steps per second: 157, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
524409/5000000: episode: 248, duration: 26.292s, episode steps: 4109, steps per second: 156, episode reward: -9.000, mean reward: -0.002 [-1.000, 1.000], mean action:
528411/5000000: episode: 249, duration: 25.524s, episode steps: 4002, steps per second: 157, episode reward: -13.000, mean reward: -0.003 [-1.000, 1.000], mean action:
532163/5000000: episode: 250, duration: 23.925s, episode steps: 3752, steps per second: 157, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
534257/5000000: episode: 251, duration: 13.391s, episode steps: 2094, steps per second: 156, episode reward: -21.000, mean reward: -0.010 [-1.000, 0.000], mean action:
539342/5000000: episode: 252, duration: 32.499s, episode steps: 5085, steps per second: 156, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean action:
544155/5000000: episode: 253, duration: 30.748s, episode steps: 4813, steps per second: 157, episode reward: -10.000, mean reward: -0.002 [-1.000, 1.000], mean action:
547734/5000000: episode: 254, duration: 23.240s, episode steps: 3579, steps per second: 154, episode reward: -14.000, mean reward: -0.004 [-1.000, 1.000], mean action:
550854/5000000: episode: 255, duration: 19.735s, episode steps: 3120, steps per second: 158, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean action:
554701/5000000: episode: 256, duration: 24.595s, episode steps: 3847, steps per second: 156, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean action:
559300/5000000: episode: 257, duration: 29.362s, episode steps: 4599, steps per second: 157, episode reward: 2.000, mean reward: 0.000 [-1.000, 1.000], mean action:
563209/5000000: episode: 258, duration: 24.906s, episode steps: 3909, steps per second: 157, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean action:
566266/5000000: episode: 259, duration: 19.383s, episode steps: 3057, steps per second: 158, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean action:
569374/5000000: episode: 260, duration: 19.796s, episode steps: 3108, steps per second: 157, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean action:
573273/5000000: episode: 261, duration: 24.764s, episode steps: 3899, steps per second: 157, episode reward: -6.000, mean reward: -0.002 [-1.000, 1.000], mean action:
577434/5000000: episode: 262, duration: 26.402s, episode steps: 4161, steps per second: 158, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean action:
582383/5000000: episode: 263, duration: 31.575s, episode steps: 4949, steps per second: 157, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean action:
587114/5000000: episode: 264, duration: 30.078s, episode steps: 4731, steps per second: 157, episode reward: 2.000, mean reward: 0.000 [-1.000, 1.000], mean action:
591054/5000000: episode: 265, duration: 25.127s, episode steps: 3940, steps per second: 157, episode reward: 9.000, mean reward: 0.002 [-1.000, 1.000], mean action:
595212/5000000: episode: 266, duration: 27.003s, episode steps: 4158, steps per second: 154, episode reward: -8.000, mean reward: -0.002 [-1.000, 1.000], mean action:
599290/5000000: episode: 267, duration: 25.848s, episode steps: 4078, steps per second: 158, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean action:
602875/5000000: episode: 268, duration: 22.737s, episode steps: 3585, steps per second: 158, episode reward: 11.000, mean reward: 0.003 [-1.000, 1.000], mean action:
607395/5000000: episode: 269, duration: 28.490s, episode steps: 4520, steps per second: 159, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean action:
611876/5000000: episode: 270, duration: 28.409s, episode steps: 4481, steps per second: 158, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean action:
615191/5000000: episode: 271, duration: 20.949s, episode steps: 3315, steps per second: 158, episode reward: -15.000, mean reward: -0.005 [-1.000, 1.000], mean action:
618289/5000000: episode: 272, duration: 19.730s, episode steps: 3098, steps per second: 157, episode reward: -19.000, mean reward: -0.006 [-1.000, 1.000], mean action:
622397/5000000: episode: 273, duration: 26.012s, episode steps: 4108, steps per second: 158, episode reward: -10.000, mean reward: -0.002 [-1.000, 1.000], mean action:
625780/5000000: episode: 274, duration: 21.419s, episode steps: 3383, steps per second: 158, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean action:
628569/5000000: episode: 275, duration: 17.698s, episode steps: 2789, steps per second: 158, episode reward: -19.000, mean reward: -0.007 [-1.000, 1.000], mean action:
632443/5000000: episode: 276, duration: 24.601s, episode steps: 3874, steps per second: 157, episode reward: -9.000, mean reward: -0.002 [-1.000, 1.000], mean action:
636705/5000000: episode: 277, duration: 27.003s, episode steps: 4262, steps per second: 158, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean action:
641018/5000000: episode: 278, duration: 27.498s, episode steps: 4313, steps per second: 157, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean action:
644843/5000000: episode: 279, duration: 25.021s, episode steps: 3825, steps per second: 153, episode reward: -14.000, mean reward: -0.004 [-1.000, 1.000], mean action:
648597/5000000: episode: 280, duration: 24.022s, episode steps: 3754, steps per second: 156, episode reward: -13.000, mean reward: -0.003 [-1.000, 1.000], mean action:
653243/5000000: episode: 281, duration: 29.299s, episode steps: 4646, steps per second: 159, episode reward: 2.000, mean reward: 0.000 [-1.000, 1.000], mean action:
658464/5000000: episode: 282, duration: 32.966s, episode steps: 5221, steps per second: 158, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean action:
663163/5000000: episode: 283, duration: 30.064s, episode steps: 4699, steps per second: 156, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
667957/5000000: episode: 284, duration: 30.991s, episode steps: 4794, steps per second: 155, episode reward: 1.000, mean reward: 0.000 [-1.000, 1.000], mean action:
672688/5000000: episode: 285, duration: 30.716s, episode steps: 4731, steps per second: 154, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
676985/5000000: episode: 286, duration: 28.234s, episode steps: 4297, steps per second: 152, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean action:
680324/5000000: episode: 287, duration: 22.007s, episode steps: 3339, steps per second: 152, episode reward: -10.000, mean reward: -0.003 [-1.000, 1.000], mean action:
684228/5000000: episode: 288, duration: 26.064s, episode steps: 3904, steps per second: 150, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean action:
687820/5000000: episode: 289, duration: 23.769s, episode steps: 3592, steps per second: 151, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
692275/5000000: episode: 290, duration: 29.823s, episode steps: 4455, steps per second: 149, episode reward: -2.000, mean reward: -0.000 [-1.000, 1.000], mean action:
696669/5000000: episode: 291, duration: 28.691s, episode steps: 4394, steps per second: 153, episode reward: -6.000, mean reward: -0.001 [-1.000, 1.000], mean action:
701270/5000000: episode: 292, duration: 29.949s, episode steps: 4601, steps per second: 154, episode reward: -2.000, mean reward: -0.000 [-1.000, 1.000], mean action:
705261/5000000: episode: 293, duration: 26.077s, episode steps: 3991, steps per second: 153, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
709796/5000000: episode: 294, duration: 29.408s, episode steps: 4535, steps per second: 154, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
714513/5000000: episode: 295, duration: 30.703s, episode steps: 4717, steps per second: 154, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean action:
719488/5000000: episode: 296, duration: 32.594s, episode steps: 4975, steps per second: 153, episode reward: -2.000, mean reward: -0.000 [-1.000, 1.000], mean action:
723419/5000000: episode: 297, duration: 25.596s, episode steps: 3931, steps per second: 154, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
727768/5000000: episode: 298, duration: 28.365s, episode steps: 4349, steps per second: 153, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
731464/5000000: episode: 299, duration: 23.900s, episode steps: 3696, steps per second: 155, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
735916/5000000: episode: 300, duration: 29.114s, episode steps: 4452, steps per second: 153, episode reward: 2.000, mean reward: 0.000 [-1.000, 1.000], mean action:
739275/5000000: episode: 301, duration: 22.140s, episode steps: 3359, steps per second: 152, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
741597/5000000: episode: 302, duration: 15.085s, episode steps: 2322, steps per second: 154, episode reward: -13.000, mean reward: -0.006 [-1.000, 1.000], mean action:
744557/5000000: episode: 303, duration: 19.161s, episode steps: 2960, steps per second: 154, episode reward: 13.000, mean reward: 0.004 [-1.000, 1.000], mean action:
747148/5000000: episode: 304, duration: 16.731s, episode steps: 2591, steps per second: 155, episode reward: 16.000, mean reward: 0.006 [-1.000, 1.000], mean action:
751514/5000000: episode: 305, duration: 28.335s, episode steps: 4366, steps per second: 154, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
755373/5000000: episode: 306, duration: 25.013s, episode steps: 3859, steps per second: 154, episode reward: -5.000, mean reward: -0.001 [-1.000, 1.000], mean action:
759509/5000000: episode: 307, duration: 26.842s, episode steps: 4136, steps per second: 154, episode reward: -7.000, mean reward: -0.002 [-1.000, 1.000], mean action:
762295/5000000: episode: 308, duration: 17.997s, episode steps: 2786, steps per second: 155, episode reward: -16.000, mean reward: -0.006 [-1.000, 1.000], mean action:
765817/5000000: episode: 309, duration: 22.858s, episode steps: 3522, steps per second: 154, episode reward: -11.000, mean reward: -0.003 [-1.000, 1.000], mean action:
769387/5000000: episode: 310, duration: 23.193s, episode steps: 3570, steps per second: 154, episode reward: -13.000, mean reward: -0.004 [-1.000, 1.000], mean action:
774379/5000000: episode: 311, duration: 32.399s, episode steps: 4992, steps per second: 154, episode reward: -6.000, mean reward: -0.001 [-1.000, 1.000], mean action:
778826/5000000: episode: 312, duration: 28.775s, episode steps: 4447, steps per second: 155, episode reward: -3.000, mean reward: -0.001 [-1.000, 1.000], mean action:
783472/5000000: episode: 313, duration: 30.185s, episode steps: 4646, steps per second: 154, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean action:
789050/5000000: episode: 314, duration: 36.351s, episode steps: 5578, steps per second: 153, episode reward: -4.000, mean reward: -0.001 [-1.000, 1.000], mean action:
793752/5000000: episode: 315, duration: 30.293s, episode steps: 4702, steps per second: 155, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
797695/5000000: episode: 316, duration: 25.329s, episode steps: 3943, steps per second: 156, episode reward: -13.000, mean reward: -0.003 [-1.000, 1.000], mean action:
802482/5000000: episode: 317, duration: 30.797s, episode steps: 4787, steps per second: 155, episode reward: -2.000, mean reward: -0.000 [-1.000, 1.000], mean action:
805616/5000000: episode: 318, duration: 20.152s, episode steps: 3134, steps per second: 156, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean action:
810080/5000000: episode: 319, duration: 28.649s, episode steps: 4464, steps per second: 156, episode reward: 6.000, mean reward: 0.001 [-1.000, 1.000], mean action:
814417/5000000: episode: 320, duration: 27.896s, episode steps: 4337, steps per second: 155, episode reward: -2.000, mean reward: -0.000 [-1.000, 1.000], mean action:
818767/5000000: episode: 321, duration: 27.807s, episode steps: 4350, steps per second: 156, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
822225/5000000: episode: 322, duration: 22.039s, episode steps: 3458, steps per second: 157, episode reward: 12.000, mean reward: 0.003 [-1.000, 1.000], mean action:
826521/5000000: episode: 323, duration: 27.602s, episode steps: 4296, steps per second: 156, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
830363/5000000: episode: 324, duration: 24.867s, episode steps: 3842, steps per second: 155, episode reward: 12.000, mean reward: 0.003 [-1.000, 1.000], mean action:
833793/5000000: episode: 325, duration: 22.395s, episode steps: 3430, steps per second: 153, episode reward: 12.000, mean reward: 0.003 [-1.000, 1.000], mean action:
837960/5000000: episode: 326, duration: 26.793s, episode steps: 4167, steps per second: 156, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:

```

| | | | | | | | | | | | | | |
|------------------|----------|------|-----------|----------|----------------|-------|-------------------|------|-----------------|----------|--------------|-------------------------|------------|
| 840525/5000000: | episode: | 327, | duration: | 16.499s, | episode steps: | 2565, | steps per second: | 155, | episode reward: | 15.000, | mean reward: | 0.006 [-1.000, 1.000], | mean actor |
| 844293/5000000: | episode: | 328, | duration: | 24.311s, | episode steps: | 3768, | steps per second: | 155, | episode reward: | -4.000, | mean reward: | -0.001 [-1.000, 1.000], | mean actor |
| 847841/5000000: | episode: | 329, | duration: | 22.789s, | episode steps: | 3548, | steps per second: | 156, | episode reward: | 6.000, | mean reward: | 0.002 [-1.000, 1.000], | mean actor |
| 852058/5000000: | episode: | 330, | duration: | 22.050s, | episode steps: | 4217, | steps per second: | 156, | episode reward: | -2.000, | mean reward: | -0.000 [-1.000, 1.000], | mean actor |
| 854918/5000000: | episode: | 331, | duration: | 18.401s, | episode steps: | 2860, | steps per second: | 155, | episode reward: | -9.000, | mean reward: | -0.003 [-1.000, 1.000], | mean actor |
| 858316/5000000: | episode: | 332, | duration: | 21.754s, | episode steps: | 3398, | steps per second: | 156, | episode reward: | 8.000, | mean reward: | 0.002 [-1.000, 1.000], | mean actor |
| 861732/5000000: | episode: | 333, | duration: | 21.869s, | episode steps: | 3416, | steps per second: | 156, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean actor |
| 864467/5000000: | episode: | 334, | duration: | 17.752s, | episode steps: | 2735, | steps per second: | 154, | episode reward: | 13.000, | mean reward: | 0.005 [-1.000, 1.000], | mean actor |
| 867990/5000000: | episode: | 335, | duration: | 22.754s, | episode steps: | 3523, | steps per second: | 155, | episode reward: | -12.000, | mean reward: | -0.003 [-1.000, 1.000], | mean actor |
| 873040/5000000: | episode: | 336, | duration: | 32.561s, | episode steps: | 5050, | steps per second: | 155, | episode reward: | -1.000, | mean reward: | -0.000 [-1.000, 1.000], | mean actor |
| 877691/5000000: | episode: | 337, | duration: | 30.161s, | episode steps: | 4651, | steps per second: | 154, | episode reward: | -1.000, | mean reward: | -0.000 [-1.000, 1.000], | mean actor |
| 882230/5000000: | episode: | 338, | duration: | 29.836s, | episode steps: | 4539, | steps per second: | 152, | episode reward: | -15.000, | mean reward: | -0.003 [-1.000, 1.000], | mean actor |
| 886712/5000000: | episode: | 339, | duration: | 29.005s, | episode steps: | 4482, | steps per second: | 155, | episode reward: | 5.000, | mean reward: | 0.001 [-1.000, 1.000], | mean actor |
| 890504/5000000: | episode: | 340, | duration: | 24.755s, | episode steps: | 3792, | steps per second: | 153, | episode reward: | 6.000, | mean reward: | 0.002 [-1.000, 1.000], | mean actor |
| 894423/5000000: | episode: | 341, | duration: | 25.747s, | episode steps: | 3919, | steps per second: | 152, | episode reward: | 9.000, | mean reward: | 0.002 [-1.000, 1.000], | mean actor |
| 898277/5000000: | episode: | 342, | duration: | 25.031s, | episode steps: | 3854, | steps per second: | 154, | episode reward: | 9.000, | mean reward: | 0.002 [-1.000, 1.000], | mean actor |
| 902292/5000000: | episode: | 343, | duration: | 26.002s, | episode steps: | 4015, | steps per second: | 154, | episode reward: | -7.000, | mean reward: | -0.002 [-1.000, 1.000], | mean actor |
| 906234/5000000: | episode: | 344, | duration: | 25.618s, | episode steps: | 3942, | steps per second: | 154, | episode reward: | 5.000, | mean reward: | 0.001 [-1.000, 1.000], | mean actor |
| 909470/5000000: | episode: | 345, | duration: | 21.067s, | episode steps: | 3236, | steps per second: | 154, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean actor |
| 912564/5000000: | episode: | 346, | duration: | 20.198s, | episode steps: | 3094, | steps per second: | 153, | episode reward: | 4.000, | mean reward: | 0.001 [-1.000, 1.000], | mean actor |
| 915572/5000000: | episode: | 347, | duration: | 19.557s, | episode steps: | 3008, | steps per second: | 154, | episode reward: | -8.000, | mean reward: | -0.003 [-1.000, 1.000], | mean actor |
| 918572/5000000: | episode: | 348, | duration: | 19.705s, | episode steps: | 3000, | steps per second: | 152, | episode reward: | 9.000, | mean reward: | 0.003 [-1.000, 1.000], | mean actor |
| 922616/5000000: | episode: | 349, | duration: | 26.354s, | episode steps: | 4044, | steps per second: | 153, | episode reward: | 3.000, | mean reward: | 0.001 [-1.000, 1.000], | mean actor |
| 926263/5000000: | episode: | 350, | duration: | 23.963s, | episode steps: | 3647, | steps per second: | 152, | episode reward: | -2.000, | mean reward: | -0.001 [-1.000, 1.000], | mean actor |
| 928791/5000000: | episode: | 351, | duration: | 16.852s, | episode steps: | 2528, | steps per second: | 150, | episode reward: | 18.000, | mean reward: | 0.007 [-1.000, 1.000], | mean actor |
| 931997/5000000: | episode: | 352, | duration: | 20.891s, | episode steps: | 3206, | steps per second: | 153, | episode reward: | 12.000, | mean reward: | 0.004 [-1.000, 1.000], | mean actor |
| 934885/5000000: | episode: | 353, | duration: | 18.918s, | episode steps: | 2888, | steps per second: | 153, | episode reward: | 17.000, | mean reward: | 0.006 [-1.000, 1.000], | mean actor |
| 939347/5000000: | episode: | 354, | duration: | 29.305s, | episode steps: | 4462, | steps per second: | 152, | episode reward: | -7.000, | mean reward: | -0.002 [-1.000, 1.000], | mean actor |
| 943577/5000000: | episode: | 355, | duration: | 27.683s, | episode steps: | 4230, | steps per second: | 153, | episode reward: | 7.000, | mean reward: | 0.002 [-1.000, 1.000], | mean actor |
| 946692/5000000: | episode: | 356, | duration: | 20.414s, | episode steps: | 3115, | steps per second: | 153, | episode reward: | 8.000, | mean reward: | 0.003 [-1.000, 1.000], | mean actor |
| 950947/5000000: | episode: | 357, | duration: | 27.817s, | episode steps: | 4255, | steps per second: | 153, | episode reward: | -3.000, | mean reward: | -0.001 [-1.000, 1.000], | mean actor |
| 954549/5000000: | episode: | 358, | duration: | 23.673s, | episode steps: | 3602, | steps per second: | 152, | episode reward: | 9.000, | mean reward: | 0.002 [-1.000, 1.000], | mean actor |
| 958989/5000000: | episode: | 359, | duration: | 28.915s, | episode steps: | 4440, | steps per second: | 154, | episode reward: | -1.000, | mean reward: | -0.000 [-1.000, 1.000], | mean actor |
| 962200/5000000: | episode: | 360, | duration: | 20.979s, | episode steps: | 3211, | steps per second: | 153, | episode reward: | -15.000, | mean reward: | -0.005 [-1.000, 1.000], | mean actor |
| 964841/5000000: | episode: | 361, | duration: | 17.343s, | episode steps: | 2641, | steps per second: | 152, | episode reward: | -15.000, | mean reward: | -0.006 [-1.000, 1.000], | mean actor |
| 968818/5000000: | episode: | 362, | duration: | 26.230s, | episode steps: | 3977, | steps per second: | 152, | episode reward: | 5.000, | mean reward: | 0.001 [-1.000, 1.000], | mean actor |
| 971043/5000000: | episode: | 363, | duration: | 14.770s, | episode steps: | 2225, | steps per second: | 151, | episode reward: | -15.000, | mean reward: | -0.007 [-1.000, 1.000], | mean actor |
| 973357/5000000: | episode: | 364, | duration: | 21.832s, | episode steps: | 3314, | steps per second: | 152, | episode reward: | -9.000, | mean reward: | -0.003 [-1.000, 1.000], | mean actor |
| 977957/5000000: | episode: | 365, | duration: | 23.973s, | episode steps: | 3600, | steps per second: | 150, | episode reward: | -13.000, | mean reward: | -0.004 [-1.000, 1.000], | mean actor |
| 981006/5000000: | episode: | 366, | duration: | 20.081s, | episode steps: | 3049, | steps per second: | 152, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean actor |
| 985104/5000000: | episode: | 367, | duration: | 26.989s, | episode steps: | 4098, | steps per second: | 152, | episode reward: | -3.000, | mean reward: | -0.001 [-1.000, 1.000], | mean actor |
| 990133/5000000: | episode: | 368, | duration: | 32.906s, | episode steps: | 5029, | steps per second: | 153, | episode reward: | 1.000, | mean reward: | 0.000 [-1.000, 1.000], | mean actor |
| 993553/5000000: | episode: | 369, | duration: | 22.366s, | episode steps: | 3420, | steps per second: | 153, | episode reward: | 12.000, | mean reward: | 0.004 [-1.000, 1.000], | mean actor |
| 996160/5000000: | episode: | 370, | duration: | 17.085s, | episode steps: | 2607, | steps per second: | 153, | episode reward: | 17.000, | mean reward: | 0.007 [-1.000, 1.000], | mean actor |
| 1000265/5000000: | episode: | 371, | duration: | 26.869s, | episode steps: | 4105, | steps per second: | 153, | episode reward: | 5.000, | mean reward: | 0.001 [-1.000, 1.000], | mean actor |
| 1002862/5000000: | episode: | 372, | duration: | 16.991s, | episode steps: | 2597, | steps per second: | 153, | episode reward: | 17.000, | mean reward: | 0.007 [-1.000, 1.000], | mean actor |
| 1007004/5000000: | episode: | 373, | duration: | 27.120s, | episode steps: | 4142, | steps per second: | 153, | episode reward: | -8.000, | mean reward: | -0.002 [-1.000, 1.000], | mean actor |
| 1010699/5000000: | episode: | 374, | duration: | 24.333s, | episode steps: | 3695, | steps per second: | 152, | episode reward: | 9.000, | mean reward: | 0.002 [-1.000, 1.000], | mean actor |
| 1014329/5000000: | episode: | 375, | duration: | 23.870s, | episode steps: | 3630, | steps per second: | 152, | episode reward: | -7.000, | mean reward: | -0.002 [-1.000, 1.000], | mean actor |
| 1019001/5000000: | episode: | 376, | duration: | 30.924s, | episode steps: | 4672, | steps per second: | 151, | episode reward: | 4.000, | mean reward: | 0.001 [-1.000, 1.000], | mean actor |
| 1023229/5000000: | episode: | 377, | duration: | 28.229s, | episode steps: | 4228, | steps per second: | 150, | episode reward: | 5.000, | mean reward: | 0.001 [-1.000, 1.000], | mean actor |
| 1028132/5000000: | episode: | 378, | duration: | 32.123s, | episode steps: | 4903, | steps per second: | 153, | episode reward: | -2.000, | mean reward: | -0.000 [-1.000, 1.000], | mean actor |
| 1031090/5000000: | episode: | 379, | duration: | 19.279s, | episode steps: | 2958, | steps per second: | 153, | episode reward: | 11.000, | mean reward: | 0.004 [-1.000, 1.000], | mean actor |
| 1035763/5000000: | episode: | 380, | duration: | 30.453s, | episode steps: | 4673, | steps per second: | 153, | episode reward: | -1.000, | mean reward: | -0.000 [-1.000, 1.000], | mean actor |
| 1039155/5000000: | episode: | 381, | duration: | 22.163s, | episode steps: | 3392, | steps per second: | 153, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean actor |
| 1043619/5000000: | episode: | 382, | duration: | 29.076s, | episode steps: | 4464, | steps per second: | 154, | episode reward: | -6.000, | mean reward: | -0.001 [-1.000, 1.000], | mean actor |
| 1047432/5000000: | episode: | 383, | duration: | 24.821s, | episode steps: | 3813, | steps per second: | 154, | episode reward: | -4.000, | mean reward: | -0.001 [-1.000, 1.000], | mean actor |
| 1051674/5000000: | episode: | 384, | duration: | 27.598s, | episode steps: | 4242, | steps per second: | 154, | episode reward: | -1.000, | mean reward: | -0.000 [-1.000, 1.000], | mean actor |
| 1054528/5000000: | episode: | 385, | duration: | 18.649s, | episode steps: | 2854, | steps per second: | 153, | episode reward: | -14.000, | mean reward: | -0.005 [-1.000, 1.000], | mean actor |
| 1057294/5000000: | episode: | 386, | duration: | 17.992s, | episode steps: | 2766, | steps per second: | 154, | episode reward: | 14.000, | mean reward: | 0.005 [-1.000, 1.000], | mean actor |
| 1059974/5000000: | episode: | 387, | duration: | 17.441s, | episode steps: | 2680, | steps per second: | 154, | episode reward: | 11.000, | mean reward: | 0.004 [-1.000, 1.000], | mean actor |
| 1063362/5000000: | episode: | 388, | duration: | 22.272s, | episode steps: | 3388, | steps per second: | 152, | episode reward: | 6.000, | mean reward: | 0.002 [-1.000, 1.000], | mean actor |
| 1066667/5000000: | episode: | 389, | duration: | 21.450s, | episode steps: | 3305, | steps per second: | 154, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean actor |
| 1069998/5000000: | episode: | 390, | duration: | 21.985s, | episode steps: | 3331, | steps per second: | 152, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean actor |
| 1074901/5000000: | episode: | 391, | duration: | 31.926s, | episode steps: | 4903, | steps per second: | 154, | episode reward: | -1.000, | mean reward: | -0.000 [-1.000, 1.000], | mean actor |
| 1077997/5000000: | episode: | 392, | duration: | 20.136s, | episode steps: | 3096, | steps per second: | 154, | episode reward: | -13.000, | mean reward: | -0.004 [-1.000, 1.000], | mean actor |
| 1081959/5000000: | episode: | 393, | duration: | 25.620s, | episode steps: | 3962, | steps per second: | 155, | episode reward: | 6.000, | mean reward: | 0.002 [-1.000, 1.000], | mean actor |
| 1084585/5000000: | episode: | 394, | duration: | 17.048s, | episode steps: | 2626, | steps per second: | 154, | episode reward: | -14.000, | mean reward: | -0.005 [-1.000, 1.000], | mean actor |
| 1089321/5000000: | episode: | 395, | duration: | 30.604s, | episode steps: | 4736, | steps per second: | 155, | episode reward: | -1.000, | mean reward: | -0.000 [-1.000, 1.000], | mean actor |
| 1094134/5000000: | episode: | 396, | duration: | 31.139s, | episode steps: | 4813, | steps per second: | 155, | episode reward: | 2.000, | mean reward: | 0.000 [-1.000, 1.000], | mean actor |
| 1098233/5000000: | episode: | 397, | duration: | 26.411s, | episode steps: | 4099, | steps per second: | 155, | episode reward: | 2.000, | mean reward: | 0.000 [-1.000, 1.000], | mean actor |
| 1101705/5000000: | episode: | 398, | duration: | 22.408s, | episode steps: | 3472, | steps per second: | 155, | episode reward: | -9.000, | mean reward: | -0.003 [-1.000, 1.000], | mean actor |
| 1104749/5000000: | episode: | 399, | duration: | 19.659s, | episode steps: | 3044, | steps per second: | 155, | episode reward: | -11.000, | mean reward: | -0.004 [-1.000, 1.000], | mean actor |
| 1108493/5000000: | episode: | 400, | duration: | 24.394s, | episode steps: | 3744, | steps per second: | 153, | episode reward: | 4.000, | mean reward: | 0.001 [-1.000, 1.000], | mean actor |
| 1112547/5000000: | episode: | 401, | duration: | 26.661s, | episode steps: | 4054, | steps per second: | 152, | episode reward: | 6.000, | mean reward: | 0.001 [-1.000, 1.000], | mean actor |
| 1115386/5000000: | episode: | 402, | duration: | 18.496s, | episode steps: | 2839, | steps per second: | 153, | episode reward: | 14.000, | mean reward: | 0.005 [-1.000, 1.000], | mean actor |
| 1118174/5000000: | episode: | 403, | duration: | 18.497s, | episode steps: | 2788, | steps per second: | 151, | episode reward: | 12.000, | mean reward: | 0.004 [-1.000, 1.000], | mean actor |
| 1121836/5000000: | episode: | 404, | duration: | 23.957s, | episode steps: | 3662, | steps per second: | 153, | episode reward: | -1.000, | mean reward: | -0.000 [-1.000, 1.000], | mean actor |
| 1125339/5000000: | episode: | 405, | duration: | 22.930s, | episode steps: | 3503, | steps per second: | 153, | episode reward: | -6.000, | mean reward: | -0.002 [-1.000, 1.000], | mean actor |
| 1128563/5000000: | episode: | 406, | duration: | 21.045s, | episode steps: | 3224, | steps per second: | 153, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean actor |
| 1131742/5000000: | episode: | 407, | duration: | 20.716s, | episode steps: | 3179, | steps per second: | 153, | episode reward: | 5.000, | mean reward: | 0.002 [-1.000, 1.000], | mean actor |
| 1134071/5000000: | episode: | 408, | duration: | 15.234s, | episode steps: | 2329, | steps per second: | 153, | episode reward: | 17.000, | mean reward: | 0.007 [-1.000, 1.000], | mean actor |
| 1136974/5000000: | episode: | 409, | duration: | 18.936s, | episode steps: | 2903, | steps per second: | 153, | episode reward: | -5.000, | mean reward: | -0.002 [-1.000, 1.000], | mean actor |
| 1140068/5000000: | episode: | 410, | duration: | 20.249s, | episode steps: | 3094, | steps per second: | 153, | episode reward: | 9.000, | mean reward: | 0.003 [-1.000, 1.000], | mean actor |
| 1142806/5000000: | episode: | 411, | duration: | 17.949s, | episode steps: | 2738, | steps per second: | 153, | episode reward: | 14.000, | mean reward: | 0.005 [-1.000, 1.000], | mean actor |
| 1146199/5000000: | episode: | 412, | duration: | 22.298s, | episode steps: | | | | | | | | |

| | | | | | | | | | | | | | |
|------------------|----------|------|-----------|----------|----------------|-------|-------------------|------|-----------------|----------|--------------|-------------------------|--------------|
| 1233157/5000000: | episode: | 436, | duration: | 23.179s, | episode steps: | 3597, | steps per second: | 155, | episode reward: | -6.000, | mean reward: | -0.002 [-1.000, 1.000], | mean action: |
| 1236497/5000000: | episode: | 437, | duration: | 21.593s, | episode steps: | 3340, | steps per second: | 155, | episode reward: | 5.000, | mean reward: | 0.001 [-1.000, 1.000], | mean action: |
| 1240022/5000000: | episode: | 438, | duration: | 23.131s, | episode steps: | 3525, | steps per second: | 152, | episode reward: | -4.000, | mean reward: | -0.001 [-1.000, 1.000], | mean action: |
| 1243928/5000000: | episode: | 439, | duration: | 25.761s, | episode steps: | 3906, | steps per second: | 152, | episode reward: | -9.000, | mean reward: | -0.002 [-1.000, 1.000], | mean action: |
| 1247417/5000000: | episode: | 440, | duration: | 23.192s, | episode steps: | 3489, | steps per second: | 150, | episode reward: | 9.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1250212/5000000: | episode: | 441, | duration: | 18.352s, | episode steps: | 2795, | steps per second: | 152, | episode reward: | 16.000, | mean reward: | 0.006 [-1.000, 1.000], | mean action: |
| 1252767/5000000: | episode: | 442, | duration: | 16.598s, | episode steps: | 2555, | steps per second: | 154, | episode reward: | 15.000, | mean reward: | 0.006 [-1.000, 1.000], | mean action: |
| 1256282/5000000: | episode: | 443, | duration: | 22.781s, | episode steps: | 3515, | steps per second: | 154, | episode reward: | 13.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 1260229/5000000: | episode: | 444, | duration: | 25.896s, | episode steps: | 3947, | steps per second: | 152, | episode reward: | 7.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 1263115/5000000: | episode: | 445, | duration: | 18.658s, | episode steps: | 2886, | steps per second: | 155, | episode reward: | 14.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 1266318/5000000: | episode: | 446, | duration: | 20.826s, | episode steps: | 3203, | steps per second: | 154, | episode reward: | 9.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1270834/5000000: | episode: | 447, | duration: | 29.668s, | episode steps: | 4516, | steps per second: | 152, | episode reward: | -3.000, | mean reward: | -0.001 [-1.000, 1.000], | mean action: |
| 1274147/5000000: | episode: | 448, | duration: | 21.630s, | episode steps: | 3313, | steps per second: | 153, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1277986/5000000: | episode: | 449, | duration: | 25.042s, | episode steps: | 3839, | steps per second: | 153, | episode reward: | 8.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 1281754/5000000: | episode: | 450, | duration: | 24.730s, | episode steps: | 3768, | steps per second: | 152, | episode reward: | 9.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 1285720/5000000: | episode: | 451, | duration: | 25.995s, | episode steps: | 3966, | steps per second: | 153, | episode reward: | 7.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 1290111/5000000: | episode: | 452, | duration: | 28.829s, | episode steps: | 4391, | steps per second: | 152, | episode reward: | 2.000, | mean reward: | 0.000 [-1.000, 1.000], | mean action: |
| 1293443/5000000: | episode: | 453, | duration: | 21.877s, | episode steps: | 3332, | steps per second: | 152, | episode reward: | 9.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1298005/5000000: | episode: | 454, | duration: | 30.211s, | episode steps: | 4562, | steps per second: | 151, | episode reward: | 2.000, | mean reward: | 0.000 [-1.000, 1.000], | mean action: |
| 1301547/5000000: | episode: | 455, | duration: | 23.212s, | episode steps: | 3542, | steps per second: | 153, | episode reward: | 12.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1304691/5000000: | episode: | 456, | duration: | 20.771s, | episode steps: | 3144, | steps per second: | 151, | episode reward: | 9.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1308807/5000000: | episode: | 457, | duration: | 27.506s, | episode steps: | 4116, | steps per second: | 150, | episode reward: | 2.000, | mean reward: | 0.000 [-1.000, 1.000], | mean action: |
| 1312586/5000000: | episode: | 458, | duration: | 24.918s, | episode steps: | 3779, | steps per second: | 152, | episode reward: | 5.000, | mean reward: | 0.001 [-1.000, 1.000], | mean action: |
| 1314995/5000000: | episode: | 459, | duration: | 15.852s, | episode steps: | 2409, | steps per second: | 152, | episode reward: | 16.000, | mean reward: | 0.007 [-1.000, 1.000], | mean action: |
| 1318795/5000000: | episode: | 460, | duration: | 25.139s, | episode steps: | 3800, | steps per second: | 151, | episode reward: | 5.000, | mean reward: | 0.001 [-1.000, 1.000], | mean action: |
| 1321483/5000000: | episode: | 461, | duration: | 17.768s, | episode steps: | 2688, | steps per second: | 151, | episode reward: | 14.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 1324780/5000000: | episode: | 462, | duration: | 21.811s, | episode steps: | 3297, | steps per second: | 151, | episode reward: | 9.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1327845/5000000: | episode: | 463, | duration: | 20.253s, | episode steps: | 3065, | steps per second: | 151, | episode reward: | -10.000, | mean reward: | -0.003 [-1.000, 1.000], | mean action: |
| 1331662/5000000: | episode: | 464, | duration: | 25.278s, | episode steps: | 3817, | steps per second: | 151, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1334768/5000000: | episode: | 465, | duration: | 20.536s, | episode steps: | 3106, | steps per second: | 151, | episode reward: | 13.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 1337551/5000000: | episode: | 466, | duration: | 18.417s, | episode steps: | 2783, | steps per second: | 151, | episode reward: | 11.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 1339873/5000000: | episode: | 467, | duration: | 15.394s, | episode steps: | 2322, | steps per second: | 151, | episode reward: | 17.000, | mean reward: | 0.007 [-1.000, 1.000], | mean action: |
| 1343408/5000000: | episode: | 468, | duration: | 23.335s, | episode steps: | 3535, | steps per second: | 151, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1346820/5000000: | episode: | 469, | duration: | 22.185s, | episode steps: | 3412, | steps per second: | 154, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1350756/5000000: | episode: | 470, | duration: | 25.543s, | episode steps: | 3936, | steps per second: | 154, | episode reward: | 5.000, | mean reward: | 0.001 [-1.000, 1.000], | mean action: |
| 1353281/5000000: | episode: | 471, | duration: | 16.737s, | episode steps: | 2525, | steps per second: | 151, | episode reward: | 12.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 1355948/5000000: | episode: | 472, | duration: | 17.325s, | episode steps: | 2667, | steps per second: | 154, | episode reward: | 12.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 1359163/5000000: | episode: | 473, | duration: | 20.836s, | episode steps: | 3215, | steps per second: | 154, | episode reward: | 8.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 1362426/5000000: | episode: | 474, | duration: | 21.129s, | episode steps: | 3263, | steps per second: | 154, | episode reward: | -1.000, | mean reward: | -0.000 [-1.000, 1.000], | mean action: |
| 1366350/5000000: | episode: | 475, | duration: | 25.419s, | episode steps: | 3924, | steps per second: | 154, | episode reward: | 1.000, | mean reward: | 0.000 [-1.000, 1.000], | mean action: |
| 1370098/5000000: | episode: | 476, | duration: | 24.462s, | episode steps: | 3748, | steps per second: | 153, | episode reward: | 7.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 1373099/5000000: | episode: | 477, | duration: | 19.522s, | episode steps: | 3001, | steps per second: | 154, | episode reward: | 13.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 1376633/5000000: | episode: | 478, | duration: | 23.066s, | episode steps: | 3534, | steps per second: | 153, | episode reward: | 13.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 1379880/5000000: | episode: | 479, | duration: | 21.078s, | episode steps: | 3247, | steps per second: | 154, | episode reward: | 11.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1382499/5000000: | episode: | 480, | duration: | 17.036s, | episode steps: | 2619, | steps per second: | 154, | episode reward: | 15.000, | mean reward: | 0.006 [-1.000, 1.000], | mean action: |
| 1385244/5000000: | episode: | 481, | duration: | 17.902s, | episode steps: | 2745, | steps per second: | 153, | episode reward: | 11.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 1387850/5000000: | episode: | 482, | duration: | 16.974s, | episode steps: | 2606, | steps per second: | 154, | episode reward: | 15.000, | mean reward: | 0.006 [-1.000, 1.000], | mean action: |
| 1391133/5000000: | episode: | 483, | duration: | 21.531s, | episode steps: | 3283, | steps per second: | 152, | episode reward: | 8.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 1394364/5000000: | episode: | 484, | duration: | 21.001s, | episode steps: | 3231, | steps per second: | 154, | episode reward: | 6.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 1397778/5000000: | episode: | 485, | duration: | 22.261s, | episode steps: | 3414, | steps per second: | 153, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1402659/5000000: | episode: | 486, | duration: | 32.026s, | episode steps: | 4881, | steps per second: | 152, | episode reward: | -6.000, | mean reward: | -0.001 [-1.000, 1.000], | mean action: |
| 1405860/5000000: | episode: | 487, | duration: | 20.711s, | episode steps: | 3201, | steps per second: | 155, | episode reward: | 13.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 1408899/5000000: | episode: | 488, | duration: | 19.770s, | episode steps: | 3039, | steps per second: | 154, | episode reward: | 13.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 1412170/5000000: | episode: | 489, | duration: | 24.779s, | episode steps: | 3811, | steps per second: | 154, | episode reward: | 7.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 1415633/5000000: | episode: | 490, | duration: | 19.023s, | episode steps: | 2923, | steps per second: | 154, | episode reward: | 14.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 1418562/5000000: | episode: | 491, | duration: | 18.924s, | episode steps: | 2929, | steps per second: | 155, | episode reward: | 13.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 1422142/5000000: | episode: | 492, | duration: | 23.243s, | episode steps: | 3580, | steps per second: | 154, | episode reward: | 9.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1426207/5000000: | episode: | 493, | duration: | 26.434s, | episode steps: | 4065, | steps per second: | 154, | episode reward: | -6.000, | mean reward: | -0.001 [-1.000, 1.000], | mean action: |
| 1430697/5000000: | episode: | 494, | duration: | 29.202s, | episode steps: | 4490, | steps per second: | 154, | episode reward: | 3.000, | mean reward: | 0.001 [-1.000, 1.000], | mean action: |
| 1434251/5000000: | episode: | 495, | duration: | 23.127s, | episode steps: | 3554, | steps per second: | 154, | episode reward: | 16.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 1437677/5000000: | episode: | 496, | duration: | 22.349s, | episode steps: | 3426, | steps per second: | 153, | episode reward: | 6.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 1441293/5000000: | episode: | 497, | duration: | 23.562s, | episode steps: | 3616, | steps per second: | 153, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1443692/5000000: | episode: | 498, | duration: | 15.668s, | episode steps: | 2399, | steps per second: | 153, | episode reward: | 15.000, | mean reward: | 0.006 [-1.000, 1.000], | mean action: |
| 1446224/5000000: | episode: | 499, | duration: | 16.510s, | episode steps: | 2532, | steps per second: | 153, | episode reward: | 16.000, | mean reward: | 0.006 [-1.000, 1.000], | mean action: |
| 1449652/5000000: | episode: | 500, | duration: | 22.598s, | episode steps: | 3428, | steps per second: | 152, | episode reward: | 9.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1451695/5000000: | episode: | 501, | duration: | 13.275s, | episode steps: | 2043, | steps per second: | 154, | episode reward: | 20.000, | mean reward: | 0.010 [-1.000, 1.000], | mean action: |
| 1455132/5000000: | episode: | 502, | duration: | 22.226s, | episode steps: | 3437, | steps per second: | 155, | episode reward: | 8.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 1458274/5000000: | episode: | 503, | duration: | 20.232s, | episode steps: | 3142, | steps per second: | 155, | episode reward: | 13.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 1461896/5000000: | episode: | 504, | duration: | 23.288s, | episode steps: | 3622, | steps per second: | 156, | episode reward: | 5.000, | mean reward: | 0.001 [-1.000, 1.000], | mean action: |
| 1465627/5000000: | episode: | 505, | duration: | 23.990s, | episode steps: | 3731, | steps per second: | 156, | episode reward: | -1.000, | mean reward: | -0.000 [-1.000, 1.000], | mean action: |
| 1468478/5000000: | episode: | 506, | duration: | 18.319s, | episode steps: | 2851, | steps per second: | 156, | episode reward: | 14.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 1472536/5000000: | episode: | 507, | duration: | 26.083s, | episode steps: | 4058, | steps per second: | 156, | episode reward: | 6.000, | mean reward: | 0.001 [-1.000, 1.000], | mean action: |
| 1476645/5000000: | episode: | 508, | duration: | 26.504s, | episode steps: | 4109, | steps per second: | 155, | episode reward: | -1.000, | mean reward: | -0.000 [-1.000, 1.000], | mean action: |
| 1479568/5000000: | episode: | 509, | duration: | 18.741s, | episode steps: | 2923, | steps per second: | 156, | episode reward: | -11.000, | mean reward: | -0.004 [-1.000, 1.000], | mean action: |
| 1481628/5000000: | episode: | 510, | duration: | 13.246s, | episode steps: | 2060, | steps per second: | 156, | episode reward: | 20.000, | mean reward: | 0.010 [-1.000, 1.000], | mean action: |
| 1485565/5000000: | episode: | 511, | duration: | 25.592s, | episode steps: | 3937, | steps per second: | 154, | episode reward: | 7.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 1488552/5000000: | episode: | 512, | duration: | 19.249s, | episode steps: | 2987, | steps per second: | 155, | episode reward: | 13.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 1492283/5000000: | episode: | 513, | duration: | 24.003s, | episode steps: | 3731, | steps per second: | 155, | episode reward: | 7.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 1496019/5000000: | episode: | 514, | duration: | 24.391s, | episode steps: | 3736, | steps per second: | 153, | episode reward: | 8.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 1498117/5000000: | episode: | 515, | duration: | 13.726s, | episode steps: | 2098, | steps per second: | 153, | episode reward: | 19.000, | mean reward: | 0.009 [-1.000, 1.000], | mean action: |
| 1501672/5000000: | episode: | 516, | duration: | 22.919s, | episode steps: | 3555, | steps per second: | 155, | episode reward: | 12.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1504625/5000000: | episode: | 517, | duration: | 19.021s, | episode steps: | 2953, | steps per second: | 155, | episode reward: | 14.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 1507564/5000000: | episode: | 518, | duration: | 19.182s, | episode steps: | 2939, | steps per second: | 153, | episode reward: | 11.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 1510544/5000000: | episode: | 519, | duration: | 19.461s, | episode steps: | 2980, | steps per second: | 153, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1513839/5000000: | episode: | 520, | duration: | 21.477s, | episode steps: | 3295, | steps per second: | 153, | episode reward: | 9.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1 | | | | | | | | | | | | | |

| | | | | | | | |
|------------------|---------------|--------------------|----------------------|------------------------|--------------------------|--------------------------------------|----------------|
| 1603989/5000000: | episode: 545, | duration: 26.966s, | episode steps: 4120, | steps per second: 153, | episode reward: 6.000, | mean reward: 0.001 [-1.000, 1.000], | mean action: |
| 1608050/5000000: | episode: 546, | duration: 26.518s, | episode steps: 4061, | steps per second: 153, | episode reward: 8.000, | mean reward: 0.002 [-1.000, 1.000], | mean action: |
| 1610630/5000000: | episode: 547, | duration: 16.902s, | episode steps: 2580, | steps per second: 153, | episode reward: 18.000, | mean reward: 0.007 [-1.000, 1.000], | mean action: |
| 1613786/5000000: | episode: 548, | duration: 20.756s, | episode steps: 3156, | steps per second: 152, | episode reward: 11.000, | mean reward: 0.003 [-1.000, 1.000], | mean action: |
| 1616105/5000000: | episode: 549, | duration: 15.348s, | episode steps: 2319, | steps per second: 151, | episode reward: 14.000, | mean reward: 0.006 [-1.000, 1.000], | mean action: |
| 1618673/5000000: | episode: 550, | duration: 16.913s, | episode steps: 2568, | steps per second: 152, | episode reward: 12.000, | mean reward: 0.005 [-1.000, 1.000], | mean action: |
| 1622007/5000000: | episode: 551, | duration: 22.099s, | episode steps: 3334, | steps per second: 151, | episode reward: 6.000, | mean reward: 0.002 [-1.000, 1.000], | mean action: |
| 1624776/5000000: | episode: 552, | duration: 18.251s, | episode steps: 2769, | steps per second: 152, | episode reward: 12.000, | mean reward: 0.004 [-1.000, 1.000], | mean action: |
| 1628247/5000000: | episode: 553, | duration: 22.858s, | episode steps: 3471, | steps per second: 152, | episode reward: 7.000, | mean reward: 0.002 [-1.000, 1.000], | mean action: |
| 1631669/5000000: | episode: 554, | duration: 22.555s, | episode steps: 3422, | steps per second: 152, | episode reward: 8.000, | mean reward: 0.002 [-1.000, 1.000], | mean action: |
| 1635266/5000000: | episode: 555, | duration: 23.383s, | episode steps: 3597, | steps per second: 154, | episode reward: 10.000, | mean reward: 0.003 [-1.000, 1.000], | mean action: |
| 1637976/5000000: | episode: 556, | duration: 17.934s, | episode steps: 2710, | steps per second: 151, | episode reward: 16.000, | mean reward: 0.006 [-1.000, 1.000], | mean action: |
| 1641217/5000000: | episode: 557, | duration: 21.099s, | episode steps: 3241, | steps per second: 154, | episode reward: 12.000, | mean reward: 0.004 [-1.000, 1.000], | mean action: |
| 1645780/5000000: | episode: 558, | duration: 29.862s, | episode steps: 4563, | steps per second: 153, | episode reward: -1.000, | mean reward: -0.000 [-1.000, 1.000], | mean action: |
| 1648267/5000000: | episode: 559, | duration: 16.276s, | episode steps: 2487, | steps per second: 153, | episode reward: 16.000, | mean reward: 0.006 [-1.000, 1.000], | mean action: |
| 1652186/5000000: | episode: 560, | duration: 25.643s, | episode steps: 3919, | steps per second: 153, | episode reward: 7.000, | mean reward: 0.002 [-1.000, 1.000], | mean action: |
| 1656369/5000000: | episode: 561, | duration: 27.130s, | episode steps: 4183, | steps per second: 154, | episode reward: 10.000, | mean reward: 0.002 [-1.000, 1.000], | mean action: |
| 1658931/5000000: | episode: 562, | duration: 16.904s, | episode steps: 2562, | steps per second: 152, | episode reward: 17.000, | mean reward: 0.007 [-1.000, 1.000], | mean action: |
| 1663289/5000000: | episode: 563, | duration: 28.455s, | episode steps: 4358, | steps per second: 153, | episode reward: 2.000, | mean reward: 0.000 [-1.000, 1.000], | mean action: |
| 1665509/5000000: | episode: 564, | duration: 14.490s, | episode steps: 2220, | steps per second: 153, | episode reward: 19.000, | mean reward: 0.009 [-1.000, 1.000], | mean action: |
| 1668023/5000000: | episode: 565, | duration: 16.657s, | episode steps: 2514, | steps per second: 151, | episode reward: 14.000, | mean reward: 0.006 [-1.000, 1.000], | mean action: |
| 1670899/5000000: | episode: 566, | duration: 18.968s, | episode steps: 2876, | steps per second: 152, | episode reward: 11.000, | mean reward: 0.004 [-1.000, 1.000], | mean action: |
| 1673877/5000000: | episode: 567, | duration: 19.524s, | episode steps: 2978, | steps per second: 153, | episode reward: 13.000, | mean reward: 0.004 [-1.000, 1.000], | mean action: |
| 1676738/5000000: | episode: 568, | duration: 18.763s, | episode steps: 2861, | steps per second: 152, | episode reward: 14.000, | mean reward: 0.005 [-1.000, 1.000], | mean action: |
| 1678788/5000000: | episode: 569, | duration: 13.474s, | episode steps: 2050, | steps per second: 152, | episode reward: 20.000, | mean reward: 0.010 [-1.000, 1.000], | mean action: |
| 1682093/5000000: | episode: 570, | duration: 21.682s, | episode steps: 3305, | steps per second: 152, | episode reward: 12.000, | mean reward: 0.004 [-1.000, 1.000], | mean action: |
| 1684354/5000000: | episode: 571, | duration: 14.875s, | episode steps: 2261, | steps per second: 152, | episode reward: 18.000, | mean reward: 0.008 [-1.000, 1.000], | mean action: |
| 1687420/5000000: | episode: 572, | duration: 20.550s, | episode steps: 3066, | steps per second: 149, | episode reward: 13.000, | mean reward: 0.004 [-1.000, 1.000], | mean action: |
| 1689946/5000000: | episode: 573, | duration: 16.707s, | episode steps: 2526, | steps per second: 151, | episode reward: 16.000, | mean reward: 0.006 [-1.000, 1.000], | mean action: |
| 1693849/5000000: | episode: 574, | duration: 25.521s, | episode steps: 3903, | steps per second: 153, | episode reward: 7.000, | mean reward: 0.002 [-1.000, 1.000], | mean action: |
| 1696367/5000000: | episode: 575, | duration: 16.417s, | episode steps: 2518, | steps per second: 153, | episode reward: 16.000, | mean reward: 0.006 [-1.000, 1.000], | mean action: |
| 1699802/5000000: | episode: 576, | duration: 22.342s, | episode steps: 3435, | steps per second: 154, | episode reward: 7.000, | mean reward: 0.002 [-1.000, 1.000], | mean action: |
| 1703273/5000000: | episode: 577, | duration: 22.388s, | episode steps: 3471, | steps per second: 155, | episode reward: 10.000, | mean reward: 0.003 [-1.000, 1.000], | mean action: |
| 1705274/5000000: | episode: 578, | duration: 12.892s, | episode steps: 2001, | steps per second: 155, | episode reward: 21.000, | mean reward: 0.010 [0.000, 1.000], | mean action: |
| 1707794/5000000: | episode: 579, | duration: 16.470s, | episode steps: 2520, | steps per second: 153, | episode reward: 13.000, | mean reward: 0.005 [-1.000, 1.000], | mean action: |
| 1710264/5000000: | episode: 580, | duration: 16.112s, | episode steps: 2470, | steps per second: 153, | episode reward: 17.000, | mean reward: 0.007 [-1.000, 1.000], | mean action: |
| 1712886/5000000: | episode: 581, | duration: 17.012s, | episode steps: 2622, | steps per second: 154, | episode reward: 13.000, | mean reward: 0.005 [-1.000, 1.000], | mean action: |
| 1715188/5000000: | episode: 582, | duration: 15.253s, | episode steps: 2302, | steps per second: 151, | episode reward: 18.000, | mean reward: 0.008 [-1.000, 1.000], | mean action: |
| 1718356/5000000: | episode: 583, | duration: 20.646s, | episode steps: 3168, | steps per second: 153, | episode reward: 9.000, | mean reward: 0.003 [-1.000, 1.000], | mean action: |
| 1720420/5000000: | episode: 584, | duration: 13.409s, | episode steps: 2064, | steps per second: 154, | episode reward: 19.000, | mean reward: 0.009 [-1.000, 1.000], | mean action: |
| 1723353/5000000: | episode: 585, | duration: 19.102s, | episode steps: 2933, | steps per second: 154, | episode reward: 12.000, | mean reward: 0.004 [-1.000, 1.000], | mean action: |
| 1726533/5000000: | episode: 586, | duration: 20.704s, | episode steps: 3180, | steps per second: 154, | episode reward: 11.000, | mean reward: 0.003 [-1.000, 1.000], | mean action: |
| 1729619/5000000: | episode: 587, | duration: 20.032s, | episode steps: 3086, | steps per second: 154, | episode reward: 15.000, | mean reward: 0.005 [-1.000, 1.000], | mean action: |
| 1732369/5000000: | episode: 588, | duration: 18.365s, | episode steps: 2750, | steps per second: 150, | episode reward: 13.000, | mean reward: 0.005 [-1.000, 1.000], | mean action: |
| 1736310/5000000: | episode: 589, | duration: 25.926s, | episode steps: 3941, | steps per second: 152, | episode reward: 2.000, | mean reward: 0.001 [-1.000, 1.000], | mean action: |
| 1738388/5000000: | episode: 590, | duration: 13.505s, | episode steps: 2078, | steps per second: 154, | episode reward: 17.000, | mean reward: 0.008 [-1.000, 1.000], | mean action: |
| 1741681/5000000: | episode: 591, | duration: 21.470s, | episode steps: 3293, | steps per second: 153, | episode reward: 9.000, | mean reward: 0.003 [-1.000, 1.000], | mean action: |
| 1744377/5000000: | episode: 592, | duration: 17.483s, | episode steps: 2696, | steps per second: 154, | episode reward: 13.000, | mean reward: 0.005 [-1.000, 1.000], | mean action: |
| 1747053/5000000: | episode: 593, | duration: 17.568s, | episode steps: 2676, | steps per second: 152, | episode reward: 11.000, | mean reward: 0.004 [-1.000, 1.000], | mean action: |
| 1750354/5000000: | episode: 594, | duration: 21.559s, | episode steps: 3301, | steps per second: 153, | episode reward: 11.000, | mean reward: 0.003 [-1.000, 1.000], | mean action: |
| 1753828/5000000: | episode: 595, | duration: 22.398s, | episode steps: 3474, | steps per second: 155, | episode reward: 6.000, | mean reward: 0.002 [-1.000, 1.000], | mean action: |
| 1757465/5000000: | episode: 596, | duration: 23.429s, | episode steps: 3637, | steps per second: 155, | episode reward: 11.000, | mean reward: 0.003 [-1.000, 1.000], | mean action: |
| 1760046/5000000: | episode: 597, | duration: 16.634s, | episode steps: 2581, | steps per second: 155, | episode reward: 18.000, | mean reward: 0.007 [-1.000, 1.000], | mean action: |
| 1763150/5000000: | episode: 598, | duration: 20.146s, | episode steps: 3104, | steps per second: 154, | episode reward: 15.000, | mean reward: 0.005 [-1.000, 1.000], | mean action: |
| 1765470/5000000: | episode: 599, | duration: 14.980s, | episode steps: 2320, | steps per second: 155, | episode reward: -13.000, | mean reward: -0.006 [-1.000, 1.000], | mean action: |
| 1768592/5000000: | episode: 600, | duration: 20.044s, | episode steps: 3122, | steps per second: 156, | episode reward: 11.000, | mean reward: 0.004 [-1.000, 1.000], | mean action: |
| 1771140/5000000: | episode: 601, | duration: 16.507s, | episode steps: 2548, | steps per second: 154, | episode reward: 14.000, | mean reward: 0.005 [-1.000, 1.000], | mean action: |
| 1773629/5000000: | episode: 602, | duration: 16.146s, | episode steps: 2489, | steps per second: 154, | episode reward: 14.000, | mean reward: 0.006 [-1.000, 1.000], | mean action: |
| 1777032/5000000: | episode: 603, | duration: 22.065s, | episode steps: 3403, | steps per second: 154, | episode reward: 3.000, | mean reward: 0.001 [-1.000, 1.000], | mean action: |
| 1779888/5000000: | episode: 604, | duration: 18.728s, | episode steps: 2856, | steps per second: 153, | episode reward: 12.000, | mean reward: 0.004 [-1.000, 1.000], | mean action: |
| 1783498/5000000: | episode: 605, | duration: 23.586s, | episode steps: 3610, | steps per second: 153, | episode reward: 6.000, | mean reward: 0.002 [-1.000, 1.000], | mean action: |
| 1786961/5000000: | episode: 606, | duration: 22.531s, | episode steps: 3463, | steps per second: 154, | episode reward: 8.000, | mean reward: 0.002 [-1.000, 1.000], | mean action: |
| 1789077/5000000: | episode: 607, | duration: 13.760s, | episode steps: 2116, | steps per second: 154, | episode reward: 17.000, | mean reward: 0.008 [-1.000, 1.000], | mean action: |
| 1792319/5000000: | episode: 608, | duration: 21.087s, | episode steps: 3242, | steps per second: 154, | episode reward: 5.000, | mean reward: 0.002 [-1.000, 1.000], | mean action: |
| 1795340/5000000: | episode: 609, | duration: 19.714s, | episode steps: 3021, | steps per second: 153, | episode reward: 9.000, | mean reward: 0.003 [-1.000, 1.000], | mean action: |
| 1798093/5000000: | episode: 610, | duration: 17.929s, | episode steps: 2753, | steps per second: 154, | episode reward: 12.000, | mean reward: 0.004 [-1.000, 1.000], | mean action: |
| 1800925/5000000: | episode: 611, | duration: 18.291s, | episode steps: 2832, | steps per second: 155, | episode reward: -3.000, | mean reward: -0.001 [-1.000, 1.000], | mean action: |
| 1804143/5000000: | episode: 612, | duration: 20.906s, | episode steps: 3218, | steps per second: 154, | episode reward: 11.000, | mean reward: 0.003 [-1.000, 1.000], | mean action: |
| 1806760/5000000: | episode: 613, | duration: 16.933s, | episode steps: 2617, | steps per second: 155, | episode reward: 17.000, | mean reward: 0.006 [-1.000, 1.000], | mean action: |
| 1810479/5000000: | episode: 614, | duration: 24.224s, | episode steps: 3719, | steps per second: 154, | episode reward: 8.000, | mean reward: 0.002 [-1.000, 1.000], | mean action: |
| 1813193/5000000: | episode: 615, | duration: 17.783s, | episode steps: 2714, | steps per second: 153, | episode reward: 14.000, | mean reward: 0.005 [-1.000, 1.000], | mean action: |
| 1815546/5000000: | episode: 616, | duration: 15.500s, | episode steps: 2353, | steps per second: 152, | episode reward: 16.000, | mean reward: 0.007 [-1.000, 1.000], | mean action: |
| 1818063/5000000: | episode: 617, | duration: 16.343s, | episode steps: 2517, | steps per second: 154, | episode reward: 14.000, | mean reward: 0.006 [-1.000, 1.000], | mean action: |
| 1821078/5000000: | episode: 618, | duration: 19.831s, | episode steps: 3015, | steps per second: 152, | episode reward: 12.000, | mean reward: 0.004 [-1.000, 1.000], | mean action: |
| 1824130/5000000: | episode: 619, | duration: 20.211s, | episode steps: 3052, | steps per second: 151, | episode reward: 9.000, | mean reward: 0.003 [-1.000, 1.000], | mean action: |
| 1826960/5000000: | episode: 620, | duration: 18.794s, | episode steps: 2830, | steps per second: 151, | episode reward: 10.000, | mean reward: 0.004 [-1.000, 1.000], | mean action: |
| 1830569/5000000: | episode: 621, | duration: 23.887s, | episode steps: 3609, | steps per second: 151, | episode reward: 3.000, | mean reward: 0.001 [-1.000, 1.000], | mean action: |
| 1833041/5000000: | episode: 622, | duration: 16.501s, | episode steps: 2472, | steps per second: 150, | episode reward: 16.000, | mean reward: 0.006 [-1.000, 1.000], | mean action: |
| 1835462/5000000: | episode: 623, | duration: 15.877s, | episode steps: 2421, | steps per second: 152, | episode reward: 15.000, | mean reward: 0.006 [-1.000, 1.000], | mean action: |
| 1838788/5000000: | episode: 624, | duration: 21.981s, | episode steps: 3326, | steps per second: 151, | episode reward: 4.000, | mean reward: 0.001 [-1.000, 1.000], | mean action: |
| 1841386/5000000: | episode: 625, | duration: 17.083s, | episode steps: 2598, | steps per second: 152, | episode reward: 10.000, | mean reward: 0.004 [-1.000, 1.000], | mean action: |
| 1843637/5000000: | episode: 626, | duration: 14.623s, | episode steps: 2251, | steps per second: 154, | episode reward: 16.000, | mean reward: 0.007 [-1.000, 1.000], | mean action: |
| 1846468/5000000: | episode: 627, | duration: 18.345s, | episode steps: 2831, | steps per second: 154, | episode reward: 10.000, | mean reward: 0.004 [-1.000, 1.000], | mean action: |
| 1848994/5000000: | episode: 628, | duration: 16.799s, | episode steps: 2526, | steps per second: 150, | episode reward: 16.000, | mean reward: 0.006 [-1.000, 1.000], | mean action: |
| 1852078/5000000: | episode: 629, | duration: 20.434s, | episode steps: 3084, | steps per second: 151, | episode reward: 7.000, | mean reward: 0.002 [-1.000, 1.000], | mean action: |
| 1855093/5000000: | episode: 630, | duration: 20.187s, | episode steps: 3015, | steps per second: 149, | episode reward: 12.000, | mean reward: 0.004 [-1.000, 1.000], | mean action: |
| 1857907/5000000: | episode: 631, | duration: 18.572s, | episode steps: 2814, | steps per second: 152, | episode reward: 17.000, | mean reward: 0.006 [-1.000, 1.000], | mean action: |
| 1861018/5000000: | episode: 632, | duration: 20.608s, | episode steps: 3111, | steps per second: 151, | episode reward: 10.000, | mean reward: 0.003 [-1.000, 1.000], | mean action: |
| 1863617/5000000: | episode: 633, | duration: 17.174s, | episode steps: 2599, | steps per second: 151, | episode reward: 13.000, | mean reward: 0.005 [-1.000, 1.000], | mean action: |
| 1865860/5000000: | episode: 634, | duration: 14.771s, | episode steps: 2243, | steps per second: 152, | episode reward: 17.000, | mean reward: 0.008 [-1.000, 1.000], | mean action: |
| 1868645/5000000: | episode: 635, | duration: 18.428s, | episode steps: 2785, | steps per second: 151, | episode reward: 7.000, | mean reward: 0.003 [-1.000, 1.000], | mean action: |
| 1870972/5000000: | episode: 636, | duration: 15.420s, | episode steps: 2327, | steps per second: 151, | episode reward: 14.000, | mean reward: 0.006 [-1.000, 1.000], | mean action: |
| 1873405/5000000: | episode: 637, | duration: 16.130s, | episode steps: 2433, | steps per second: 151, | episode reward: 15.000, | mean reward: 0.006 [-1.000, 1.000], | mean action: |
| 1875437/5000000: | episode: 638, | duration: 13.725s, | episode steps: 2032, | steps per second: 148, | episode reward: 18.000, | mean reward: 0.009 [-1.000, 1.000], | mean action: |
| 1878803/5000000: | episode: 639, | duration: 22.159s, | episode steps: 3366, | steps per second: 152, | episode reward: 5.000, | mean reward: 0.001 [-1.000, 1.000], | mean action: |
| 1881329/5000000: | episode: 640, | duration: 16.570s, | episode steps: 2526, | steps per second: 152, | episode reward: 12.000, | mean reward: 0.005 [-1.000, 1.000], | mean action: |
| 1883919/5000000: | episode: 641, | duration: 17.040s, | episode steps: 2590, | steps per second: 152, | episode reward: 12.000, | mean reward: 0.005 [-1.000, 1.000], | mean action:</ |

| | | | | | | | | | | | | | |
|------------------|----------|------|-----------|----------|----------------|-------|-------------------|------|-----------------|---------|--------------|-------------------------|--------------|
| 1923056/5000000: | episode: | 654, | duration: | 17.208s, | episode steps: | 2595, | steps per second: | 151, | episode reward: | 14.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 1925800/5000000: | episode: | 655, | duration: | 17.694s, | episode steps: | 2744, | steps per second: | 155, | episode reward: | 15.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 1929475/5000000: | episode: | 656, | duration: | 23.732s, | episode steps: | 3675, | steps per second: | 155, | episode reward: | 5.000, | mean reward: | 0.001 [-1.000, 1.000], | mean action: |
| 1933924/5000000: | episode: | 657, | duration: | 28.934s, | episode steps: | 4449, | steps per second: | 154, | episode reward: | -1.000, | mean reward: | -0.000 [-1.000, 1.000], | mean action: |
| 1936962/5000000: | episode: | 658, | duration: | 19.724s, | episode steps: | 3038, | steps per second: | 154, | episode reward: | 8.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1940913/5000000: | episode: | 659, | duration: | 25.335s, | episode steps: | 3951, | steps per second: | 156, | episode reward: | 3.000, | mean reward: | 0.001 [-1.000, 1.000], | mean action: |
| 1943706/5000000: | episode: | 660, | duration: | 17.943s, | episode steps: | 2793, | steps per second: | 156, | episode reward: | 14.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 1947162/5000000: | episode: | 661, | duration: | 22.378s, | episode steps: | 3456, | steps per second: | 154, | episode reward: | 11.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1950532/5000000: | episode: | 662, | duration: | 21.870s, | episode steps: | 3370, | steps per second: | 154, | episode reward: | 11.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1954185/5000000: | episode: | 663, | duration: | 23.454s, | episode steps: | 3653, | steps per second: | 156, | episode reward: | 1.000, | mean reward: | 0.000 [-1.000, 1.000], | mean action: |
| 1957013/5000000: | episode: | 664, | duration: | 18.233s, | episode steps: | 2828, | steps per second: | 155, | episode reward: | 6.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 1959905/5000000: | episode: | 665, | duration: | 18.469s, | episode steps: | 2892, | steps per second: | 157, | episode reward: | 16.000, | mean reward: | 0.006 [-1.000, 1.000], | mean action: |
| 1964035/5000000: | episode: | 666, | duration: | 26.385s, | episode steps: | 4130, | steps per second: | 157, | episode reward: | 2.000, | mean reward: | 0.000 [-1.000, 1.000], | mean action: |
| 1966763/5000000: | episode: | 667, | duration: | 17.508s, | episode steps: | 2728, | steps per second: | 156, | episode reward: | 12.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 1970074/5000000: | episode: | 668, | duration: | 21.612s, | episode steps: | 3311, | steps per second: | 153, | episode reward: | 9.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1973175/5000000: | episode: | 669, | duration: | 19.707s, | episode steps: | 3101, | steps per second: | 157, | episode reward: | 15.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 1976117/5000000: | episode: | 670, | duration: | 18.880s, | episode steps: | 2942, | steps per second: | 156, | episode reward: | 13.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 1978318/5000000: | episode: | 671, | duration: | 14.154s, | episode steps: | 2201, | steps per second: | 156, | episode reward: | 18.000, | mean reward: | 0.008 [-1.000, 1.000], | mean action: |
| 1981539/5000000: | episode: | 672, | duration: | 20.780s, | episode steps: | 3221, | steps per second: | 155, | episode reward: | 8.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 1984811/5000000: | episode: | 673, | duration: | 21.044s, | episode steps: | 3272, | steps per second: | 155, | episode reward: | 11.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 1987866/5000000: | episode: | 674, | duration: | 19.684s, | episode steps: | 3055, | steps per second: | 155, | episode reward: | 12.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 1991888/5000000: | episode: | 675, | duration: | 25.799s, | episode steps: | 4022, | steps per second: | 156, | episode reward: | -7.000, | mean reward: | -0.002 [-1.000, 1.000], | mean action: |
| 1994990/5000000: | episode: | 676, | duration: | 20.158s, | episode steps: | 3102, | steps per second: | 154, | episode reward: | 12.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 1997423/5000000: | episode: | 677, | duration: | 15.624s, | episode steps: | 2433, | steps per second: | 156, | episode reward: | 17.000, | mean reward: | 0.007 [-1.000, 1.000], | mean action: |
| 2000703/5000000: | episode: | 678, | duration: | 21.097s, | episode steps: | 3280, | steps per second: | 155, | episode reward: | 11.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 2003344/5000000: | episode: | 679, | duration: | 17.039s, | episode steps: | 2641, | steps per second: | 155, | episode reward: | 16.000, | mean reward: | 0.006 [-1.000, 1.000], | mean action: |
| 2006658/5000000: | episode: | 680, | duration: | 21.331s, | episode steps: | 3314, | steps per second: | 155, | episode reward: | -6.000, | mean reward: | -0.002 [-1.000, 1.000], | mean action: |
| 2011271/5000000: | episode: | 681, | duration: | 29.566s, | episode steps: | 4613, | steps per second: | 156, | episode reward: | 1.000, | mean reward: | 0.000 [-1.000, 1.000], | mean action: |
| 2015345/5000000: | episode: | 682, | duration: | 26.351s, | episode steps: | 4074, | steps per second: | 155, | episode reward: | 3.000, | mean reward: | 0.001 [-1.000, 1.000], | mean action: |
| 2019065/5000000: | episode: | 683, | duration: | 24.552s, | episode steps: | 3720, | steps per second: | 152, | episode reward: | 9.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 2022574/5000000: | episode: | 684, | duration: | 22.749s, | episode steps: | 3509, | steps per second: | 154, | episode reward: | 8.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 2025553/5000000: | episode: | 685, | duration: | 19.439s, | episode steps: | 2979, | steps per second: | 153, | episode reward: | 14.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 2028973/5000000: | episode: | 686, | duration: | 22.237s, | episode steps: | 3420, | steps per second: | 154, | episode reward: | 7.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 2033663/5000000: | episode: | 687, | duration: | 30.538s, | episode steps: | 4690, | steps per second: | 154, | episode reward: | -1.000, | mean reward: | -0.000 [-1.000, 1.000], | mean action: |
| 2036284/5000000: | episode: | 688, | duration: | 17.068s, | episode steps: | 2621, | steps per second: | 154, | episode reward: | 17.000, | mean reward: | 0.006 [-1.000, 1.000], | mean action: |
| 2040314/5000000: | episode: | 689, | duration: | 26.217s, | episode steps: | 4030, | steps per second: | 154, | episode reward: | 8.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 2042488/5000000: | episode: | 690, | duration: | 14.344s, | episode steps: | 2174, | steps per second: | 152, | episode reward: | 20.000, | mean reward: | 0.009 [-1.000, 1.000], | mean action: |
| 2044768/5000000: | episode: | 691, | duration: | 14.837s, | episode steps: | 2280, | steps per second: | 154, | episode reward: | 17.000, | mean reward: | 0.007 [-1.000, 1.000], | mean action: |
| 2048059/5000000: | episode: | 692, | duration: | 21.226s, | episode steps: | 3291, | steps per second: | 155, | episode reward: | 7.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 2051526/5000000: | episode: | 693, | duration: | 22.350s, | episode steps: | 3467, | steps per second: | 155, | episode reward: | 6.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 2053724/5000000: | episode: | 694, | duration: | 14.209s, | episode steps: | 2198, | steps per second: | 155, | episode reward: | 19.000, | mean reward: | 0.009 [-1.000, 1.000], | mean action: |
| 2056600/5000000: | episode: | 695, | duration: | 18.547s, | episode steps: | 2876, | steps per second: | 155, | episode reward: | 11.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 2059856/5000000: | episode: | 696, | duration: | 21.021s, | episode steps: | 3256, | steps per second: | 155, | episode reward: | 11.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 2063502/5000000: | episode: | 697, | duration: | 23.576s, | episode steps: | 3646, | steps per second: | 155, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 2066342/5000000: | episode: | 698, | duration: | 18.622s, | episode steps: | 2840, | steps per second: | 153, | episode reward: | 15.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 2068950/5000000: | episode: | 699, | duration: | 16.783s, | episode steps: | 2608, | steps per second: | 155, | episode reward: | 14.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 2071470/5000000: | episode: | 700, | duration: | 16.243s, | episode steps: | 2520, | steps per second: | 155, | episode reward: | 17.000, | mean reward: | 0.007 [-1.000, 1.000], | mean action: |
| 2075048/5000000: | episode: | 701, | duration: | 23.152s, | episode steps: | 3578, | steps per second: | 155, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 2078036/5000000: | episode: | 702, | duration: | 19.246s, | episode steps: | 2988, | steps per second: | 155, | episode reward: | 13.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 2080225/5000000: | episode: | 703, | duration: | 14.089s, | episode steps: | 2189, | steps per second: | 155, | episode reward: | 19.000, | mean reward: | 0.009 [-1.000, 1.000], | mean action: |
| 2082445/5000000: | episode: | 704, | duration: | 14.288s, | episode steps: | 2220, | steps per second: | 155, | episode reward: | 19.000, | mean reward: | 0.009 [-1.000, 1.000], | mean action: |
| 2084996/5000000: | episode: | 705, | duration: | 16.425s, | episode steps: | 2551, | steps per second: | 155, | episode reward: | 14.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 2087782/5000000: | episode: | 706, | duration: | 17.904s, | episode steps: | 2786, | steps per second: | 156, | episode reward: | 15.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 2090104/5000000: | episode: | 707, | duration: | 21.051s, | episode steps: | 3232, | steps per second: | 154, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 2094701/5000000: | episode: | 708, | duration: | 23.810s, | episode steps: | 3687, | steps per second: | 155, | episode reward: | 3.000, | mean reward: | 0.001 [-1.000, 1.000], | mean action: |
| 2097407/5000000: | episode: | 709, | duration: | 17.369s, | episode steps: | 2706, | steps per second: | 156, | episode reward: | 14.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 2099849/5000000: | episode: | 710, | duration: | 15.950s, | episode steps: | 2442, | steps per second: | 153, | episode reward: | 15.000, | mean reward: | 0.006 [-1.000, 1.000], | mean action: |
| 2102882/5000000: | episode: | 711, | duration: | 19.757s, | episode steps: | 3033, | steps per second: | 154, | episode reward: | 11.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 2106483/5000000: | episode: | 712, | duration: | 23.565s, | episode steps: | 3601, | steps per second: | 153, | episode reward: | 4.000, | mean reward: | 0.001 [-1.000, 1.000], | mean action: |
| 2109899/5000000: | episode: | 713, | duration: | 22.442s, | episode steps: | 3416, | steps per second: | 152, | episode reward: | 8.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 2113195/5000000: | episode: | 714, | duration: | 21.692s, | episode steps: | 3296, | steps per second: | 152, | episode reward: | 11.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 2116694/5000000: | episode: | 715, | duration: | 22.543s, | episode steps: | 3499, | steps per second: | 155, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 2118960/5000000: | episode: | 716, | duration: | 14.615s, | episode steps: | 2266, | steps per second: | 155, | episode reward: | 17.000, | mean reward: | 0.008 [-1.000, 1.000], | mean action: |
| 2121273/5000000: | episode: | 717, | duration: | 14.873s, | episode steps: | 2313, | steps per second: | 156, | episode reward: | 17.000, | mean reward: | 0.007 [-1.000, 1.000], | mean action: |
| 2124492/5000000: | episode: | 718, | duration: | 21.027s, | episode steps: | 3219, | steps per second: | 153, | episode reward: | 13.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 2126715/5000000: | episode: | 719, | duration: | 14.517s, | episode steps: | 2223, | steps per second: | 153, | episode reward: | 17.000, | mean reward: | 0.008 [-1.000, 1.000], | mean action: |
| 2129012/5000000: | episode: | 720, | duration: | 14.915s, | episode steps: | 2297, | steps per second: | 154, | episode reward: | 16.000, | mean reward: | 0.007 [-1.000, 1.000], | mean action: |
| 2131937/5000000: | episode: | 721, | duration: | 19.247s, | episode steps: | 2925, | steps per second: | 152, | episode reward: | 14.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 2135237/5000000: | episode: | 722, | duration: | 22.229s, | episode steps: | 3300, | steps per second: | 148, | episode reward: | 9.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 2138103/5000000: | episode: | 723, | duration: | 19.212s, | episode steps: | 2866, | steps per second: | 149, | episode reward: | 13.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 2141658/5000000: | episode: | 724, | duration: | 23.676s, | episode steps: | 3555, | steps per second: | 150, | episode reward: | 6.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 2145813/5000000: | episode: | 725, | duration: | 27.476s, | episode steps: | 4155, | steps per second: | 151, | episode reward: | 4.000, | mean reward: | 0.001 [-1.000, 1.000], | mean action: |
| 2148353/5000000: | episode: | 726, | duration: | 16.517s, | episode steps: | 2540, | steps per second: | 154, | episode reward: | 12.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 2150889/5000000: | episode: | 727, | duration: | 16.494s, | episode steps: | 2536, | steps per second: | 154, | episode reward: | 16.000, | mean reward: | 0.006 [-1.000, 1.000], | mean action: |
| 2153395/5000000: | episode: | 728, | duration: | 16.515s, | episode steps: | 2506, | steps per second: | 152, | episode reward: | 15.000, | mean reward: | 0.006 [-1.000, 1.000], | mean action: |
| 2156618/5000000: | episode: | 729, | duration: | 20.852s, | episode steps: | 3223, | steps per second: | 155, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 2159615/5000000: | episode: | 730, | duration: | 20.049s, | episode steps: | 2997, | steps per second: | 149, | episode reward: | 13.000, | mean reward: | 0.004 [-1.000, 1.000], | mean action: |
| 2162384/5000000: | episode: | 731, | duration: | 18.135s, | episode steps: | 2769, | steps per second: | 153, | episode reward: | 15.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 2165817/5000000: | episode: | 732, | duration: | 22.136s, | episode steps: | 3433, | steps per second: | 155, | episode reward: | 8.000, | mean reward: | 0.002 [-1.000, 1.000], | mean action: |
| 2169807/5000000: | episode: | 733, | duration: | 25.709s, | episode steps: | 3990, | steps per second: | 155, | episode reward: | 4.000, | mean reward: | 0.001 [-1.000, 1.000], | mean action: |
| 2172613/5000000: | episode: | 734, | duration: | 18.043s, | episode steps: | 2806, | steps per second: | 156, | episode reward: | 14.000, | mean reward: | 0.005 [-1.000, 1.000], | mean action: |
| 2176913/5000000: | episode: | 735, | duration: | 27.526s, | episode steps: | 4300, | steps per second: | 156, | episode reward: | 1.000, | mean reward: | 0.000 [-1.000, 1.000], | mean action: |
| 2180421/5000000: | episode: | 736, | duration: | 22.398s, | episode steps: | 3508, | steps per second: | 157, | episode reward: | 9.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 2184097/5000000: | episode: | 737, | duration: | 23.829s, | episode steps: | 3676, | steps per second: | 154, | episode reward: | 10.000, | mean reward: | 0.003 [-1.000, 1.000], | mean action: |
| 2187464/5000000: | episode: | 738, | duration: | 21.503s, | episode steps: | 3367, | steps per second: | 157, | episode reward: | 8.000, | mean reward: | 0.002 [-1.000, 1.000], | mean |

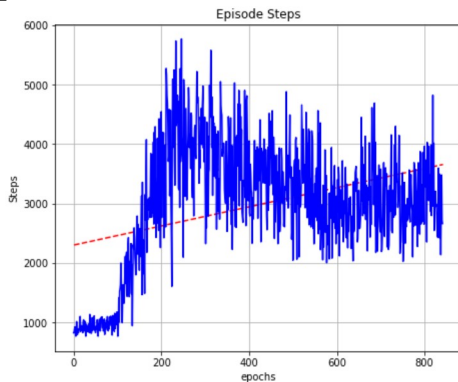
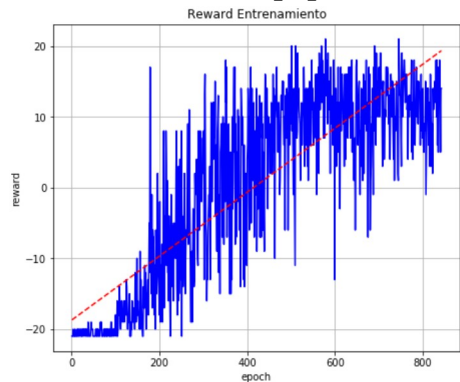
```

2260081/5000000: episode: 763, duration: 17.659s, episode steps: 2733, steps per second: 155, episode reward: 16.000, mean reward: 0.006 [-1.000, 1.000], mean actor
2262960/5000000: episode: 764, duration: 18.581s, episode steps: 2879, steps per second: 155, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2265868/5000000: episode: 765, duration: 18.744s, episode steps: 2908, steps per second: 155, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2269432/5000000: episode: 766, duration: 22.988s, episode steps: 3564, steps per second: 155, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
2272300/5000000: episode: 767, duration: 18.488s, episode steps: 2868, steps per second: 155, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2274987/5000000: episode: 768, duration: 17.425s, episode steps: 2687, steps per second: 154, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean actor
2278285/5000000: episode: 769, duration: 21.508s, episode steps: 3298, steps per second: 153, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actor
2281771/5000000: episode: 770, duration: 22.461s, episode steps: 3486, steps per second: 155, episode reward: 13.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2284997/5000000: episode: 771, duration: 20.813s, episode steps: 3226, steps per second: 155, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2287932/5000000: episode: 772, duration: 18.990s, episode steps: 2935, steps per second: 155, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2290813/5000000: episode: 773, duration: 18.643s, episode steps: 2881, steps per second: 155, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean actor
2292912/5000000: episode: 774, duration: 13.693s, episode steps: 2099, steps per second: 153, episode reward: 18.000, mean reward: 0.009 [-1.000, 1.000], mean actor
2295873/5000000: episode: 775, duration: 19.141s, episode steps: 2961, steps per second: 155, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2298492/5000000: episode: 776, duration: 16.960s, episode steps: 2619, steps per second: 154, episode reward: 12.000, mean reward: 0.005 [-1.000, 1.000], mean actor
2302179/5000000: episode: 777, duration: 23.841s, episode steps: 3687, steps per second: 155, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
2305772/5000000: episode: 778, duration: 23.557s, episode steps: 3593, steps per second: 153, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2308729/5000000: episode: 779, duration: 19.084s, episode steps: 2957, steps per second: 155, episode reward: 13.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2311361/5000000: episode: 780, duration: 17.070s, episode steps: 2632, steps per second: 154, episode reward: 18.000, mean reward: 0.007 [-1.000, 1.000], mean actor
2314931/5000000: episode: 781, duration: 23.090s, episode steps: 3570, steps per second: 155, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2318308/5000000: episode: 782, duration: 21.794s, episode steps: 3377, steps per second: 155, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2321465/5000000: episode: 783, duration: 20.368s, episode steps: 3157, steps per second: 155, episode reward: 11.000, mean reward: 0.003 [-1.000, 1.000], mean actor
2324650/5000000: episode: 784, duration: 20.767s, episode steps: 3185, steps per second: 153, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2326295/5000000: episode: 785, duration: 14.731s, episode steps: 2275, steps per second: 154, episode reward: 17.000, mean reward: 0.007 [-1.000, 1.000], mean actor
2329959/5000000: episode: 786, duration: 19.472s, episode steps: 3034, steps per second: 156, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actor
2333645/5000000: episode: 787, duration: 23.823s, episode steps: 3686, steps per second: 155, episode reward: 13.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2336419/5000000: episode: 788, duration: 17.806s, episode steps: 2774, steps per second: 156, episode reward: 15.000, mean reward: 0.005 [-1.000, 1.000], mean actor
2339737/5000000: episode: 789, duration: 21.370s, episode steps: 3318, steps per second: 155, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
2342221/5000000: episode: 790, duration: 15.924s, episode steps: 2484, steps per second: 156, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean actor
2346094/5000000: episode: 791, duration: 24.836s, episode steps: 3873, steps per second: 156, episode reward: 14.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2349434/5000000: episode: 792, duration: 21.431s, episode steps: 3340, steps per second: 156, episode reward: 13.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2352087/5000000: episode: 793, duration: 17.473s, episode steps: 2653, steps per second: 152, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean actor
2355346/5000000: episode: 794, duration: 20.919s, episode steps: 3259, steps per second: 156, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2358663/5000000: episode: 795, duration: 21.272s, episode steps: 3317, steps per second: 156, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2361358/5000000: episode: 796, duration: 17.295s, episode steps: 2695, steps per second: 156, episode reward: 16.000, mean reward: 0.006 [-1.000, 1.000], mean actor
2365214/5000000: episode: 797, duration: 24.653s, episode steps: 3856, steps per second: 156, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2367935/5000000: episode: 798, duration: 17.452s, episode steps: 2721, steps per second: 156, episode reward: 13.000, mean reward: 0.005 [-1.000, 1.000], mean actor
2371006/5000000: episode: 799, duration: 19.912s, episode steps: 3071, steps per second: 154, episode reward: 7.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2373656/5000000: episode: 800, duration: 16.950s, episode steps: 2650, steps per second: 156, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean actor
2377619/5000000: episode: 801, duration: 25.484s, episode steps: 3963, steps per second: 156, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2380951/5000000: episode: 802, duration: 21.341s, episode steps: 3332, steps per second: 156, episode reward: 5.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2384233/5000000: episode: 803, duration: 21.010s, episode steps: 3282, steps per second: 156, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
2387650/5000000: episode: 804, duration: 21.774s, episode steps: 3417, steps per second: 157, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
2391383/5000000: episode: 805, duration: 23.791s, episode steps: 3733, steps per second: 157, episode reward: 8.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2393734/5000000: episode: 806, duration: 15.025s, episode steps: 2351, steps per second: 156, episode reward: 15.000, mean reward: 0.006 [-1.000, 1.000], mean actor
2397761/5000000: episode: 807, duration: 25.672s, episode steps: 4027, steps per second: 157, episode reward: -1.000, mean reward: -0.000 [-1.000, 1.000], mean actio
2401105/5000000: episode: 808, duration: 21.715s, episode steps: 3344, steps per second: 154, episode reward: 4.000, mean reward: 0.001 [-1.000, 1.000], mean action:
2404618/5000000: episode: 809, duration: 22.352s, episode steps: 3513, steps per second: 157, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
2407672/5000000: episode: 810, duration: 19.514s, episode steps: 3054, steps per second: 157, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean actor
2411636/5000000: episode: 811, duration: 25.397s, episode steps: 3964, steps per second: 156, episode reward: 11.000, mean reward: 0.003 [-1.000, 1.000], mean actor
2414680/5000000: episode: 812, duration: 19.440s, episode steps: 3044, steps per second: 157, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean actor
2418154/5000000: episode: 813, duration: 22.343s, episode steps: 3474, steps per second: 155, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2421485/5000000: episode: 814, duration: 21.345s, episode steps: 3331, steps per second: 156, episode reward: 11.000, mean reward: 0.003 [-1.000, 1.000], mean actor
2424345/5000000: episode: 815, duration: 18.354s, episode steps: 2860, steps per second: 156, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2428330/5000000: episode: 816, duration: 25.478s, episode steps: 3985, steps per second: 156, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
2431124/5000000: episode: 817, duration: 17.891s, episode steps: 2794, steps per second: 156, episode reward: 16.000, mean reward: 0.006 [-1.000, 1.000], mean actor
2434610/5000000: episode: 818, duration: 22.214s, episode steps: 3486, steps per second: 157, episode reward: 10.000, mean reward: 0.003 [-1.000, 1.000], mean actor
2437881/5000000: episode: 819, duration: 20.831s, episode steps: 3271, steps per second: 157, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
2442703/5000000: episode: 820, duration: 31.080s, episode steps: 4822, steps per second: 155, episode reward: 3.000, mean reward: 0.001 [-1.000, 1.000], mean action:
2445657/5000000: episode: 821, duration: 18.877s, episode steps: 2954, steps per second: 156, episode reward: 9.000, mean reward: 0.003 [-1.000, 1.000], mean action:
2449671/5000000: episode: 822, duration: 25.988s, episode steps: 4014, steps per second: 154, episode reward: 2.000, mean reward: 0.000 [-1.000, 1.000], mean action:
2452216/5000000: episode: 823, duration: 16.285s, episode steps: 2545, steps per second: 156, episode reward: 14.000, mean reward: 0.006 [-1.000, 1.000], mean actor
2455223/5000000: episode: 824, duration: 19.267s, episode steps: 3007, steps per second: 156, episode reward: 11.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2458312/5000000: episode: 825, duration: 19.776s, episode steps: 3089, steps per second: 156, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2461518/5000000: episode: 826, duration: 20.608s, episode steps: 3206, steps per second: 156, episode reward: 13.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2464467/5000000: episode: 827, duration: 18.938s, episode steps: 2949, steps per second: 156, episode reward: 15.000, mean reward: 0.005 [-1.000, 1.000], mean actor
2467453/5000000: episode: 828, duration: 19.500s, episode steps: 2986, steps per second: 153, episode reward: 13.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2470181/5000000: episode: 829, duration: 17.579s, episode steps: 2728, steps per second: 155, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2472618/5000000: episode: 830, duration: 15.676s, episode steps: 2437, steps per second: 155, episode reward: 18.000, mean reward: 0.007 [-1.000, 1.000], mean actor
2475467/5000000: episode: 831, duration: 18.368s, episode steps: 2849, steps per second: 155, episode reward: 10.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2477894/5000000: episode: 832, duration: 15.516s, episode steps: 2427, steps per second: 156, episode reward: 18.000, mean reward: 0.007 [-1.000, 1.000], mean actor
2481490/5000000: episode: 833, duration: 23.193s, episode steps: 3596, steps per second: 155, episode reward: 6.000, mean reward: 0.002 [-1.000, 1.000], mean action:
2484241/5000000: episode: 834, duration: 17.750s, episode steps: 2751, steps per second: 155, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean actor
2487737/5000000: episode: 835, duration: 22.817s, episode steps: 3496, steps per second: 153, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action:
2490151/5000000: episode: 836, duration: 15.612s, episode steps: 2414, steps per second: 155, episode reward: 17.000, mean reward: 0.007 [-1.000, 1.000], mean actor
2493213/5000000: episode: 837, duration: 19.985s, episode steps: 3062, steps per second: 153, episode reward: 13.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2495348/5000000: episode: 838, duration: 14.132s, episode steps: 2135, steps per second: 151, episode reward: 18.000, mean reward: 0.008 [-1.000, 1.000], mean actor
2498268/5000000: episode: 839, duration: 19.289s, episode steps: 2920, steps per second: 151, episode reward: 12.000, mean reward: 0.004 [-1.000, 1.000], mean actor
2501749/5000000: episode: 840, duration: 22.586s, episode steps: 3481, steps per second: 154, episode reward: 5.000, mean reward: 0.001 [-1.000, 1.000], mean action:
2504465/5000000: episode: 841, duration: 17.692s, episode steps: 2716, steps per second: 154, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean actor
2507126/5000000: episode: 842, duration: 17.396s, episode steps: 2661, steps per second: 153, episode reward: 14.000, mean reward: 0.005 [-1.000, 1.000], mean actor

```

done, took 16297.762 seconds

Pesos guardados en: 202001121351_dqn_PongDeterministic-v0_weights.h5f



Testing for 10 episodes ...

Episode 1: reward: 11.000, steps: 3512

Episode 2: reward: 12.000, steps: 3135

Episode 3: reward: 6.000, steps: 3758

Episode 4: reward: 12.000, steps: 2784

Episode 5: reward: 16.000, steps: 2502


```
Episode 6: reward: 4.000, steps: 4062
Episode 7: reward: 15.000, steps: 2799
Episode 8: reward: 14.000, steps: 2856
Episode 9: reward: 15.000, steps: 2508
Episode 10: reward: 9.000, steps: 2855
mean_reward: 11.4
mean_steps: 3077.1
(11.4, 3077.1)
```