

```
Training for 500000 steps ...
```

1886/500000:	episode: 1,	duration: 3.933s,	episode steps: 972,	steps per second: 247,	episode reward: -21.000,	mean reward: -0.022	[-1.000, 0.000],	mean action: 2.
1886/500000:	episode: 2,	duration: 2.819s,	episode steps: 914,	steps per second: 324,	episode reward: -21.000,	mean reward: -0.023	[-1.000, 0.000],	mean action: 2.
2650/500000:	episode: 3,	duration: 2.384s,	episode steps: 764,	steps per second: 320,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action: 2.
3492/500000:	episode: 4,	duration: 2.672s,	episode steps: 842,	steps per second: 315,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action: 2.
4376/500000:	episode: 5,	duration: 2.824s,	episode steps: 884,	steps per second: 313,	episode reward: -21.000,	mean reward: -0.024	[-1.000, 0.000],	mean action: 2.
5351/500000:	episode: 6,	duration: 3.149s,	episode steps: 975,	steps per second: 310,	episode reward: -21.000,	mean reward: -0.022	[-1.000, 0.000],	mean action: 2.
6282/500000:	episode: 7,	duration: 2.823s,	episode steps: 931,	steps per second: 330,	episode reward: -21.000,	mean reward: -0.023	[-1.000, 0.000],	mean action: 2.
7279/500000:	episode: 8,	duration: 3.032s,	episode steps: 997,	steps per second: 329,	episode reward: -19.000,	mean reward: -0.019	[-1.000, 1.000],	mean action: 2.
8317/500000:	episode: 9,	duration: 3.129s,	episode steps: 1038,	steps per second: 332,	episode reward: -20.000,	mean reward: -0.019	[-1.000, 1.000],	mean action: 2.
9251/500000:	episode: 10,	duration: 2.822s,	episode steps: 934,	steps per second: 331,	episode reward: -21.000,	mean reward: -0.022	[-1.000, 0.000],	mean action: 2.
10135/500000:	episode: 11,	duration: 5.712s,	episode steps: 884,	steps per second: 155,	episode reward: -21.000,	mean reward: -0.024	[-1.000, 0.000],	mean action: 2.
11050/500000:	episode: 12,	duration: 16.478s,	episode steps: 915,	steps per second: 56,	episode reward: -20.000,	mean reward: -0.022	[-1.000, 1.000],	mean action: 2.
12027/500000:	episode: 13,	duration: 17.601s,	episode steps: 977,	steps per second: 56,	episode reward: -19.000,	mean reward: -0.019	[-1.000, 1.000],	mean action: 2.
12907/500000:	episode: 14,	duration: 15.671s,	episode steps: 880,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.024	[-1.000, 0.000],	mean action: 2.
13699/500000:	episode: 15,	duration: 14.130s,	episode steps: 792,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action: 2.
14718/500000:	episode: 16,	duration: 18.072s,	episode steps: 1019,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.021	[-1.000, 0.000],	mean action: 2.
15789/500000:	episode: 17,	duration: 18.823s,	episode steps: 1071,	steps per second: 57,	episode reward: -19.000,	mean reward: -0.018	[-1.000, 1.000],	mean action: 2.
16631/500000:	episode: 18,	duration: 14.950s,	episode steps: 842,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action: 2.
17474/500000:	episode: 19,	duration: 14.965s,	episode steps: 843,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action: 2.
18477/500000:	episode: 20,	duration: 17.505s,	episode steps: 1003,	steps per second: 57,	episode reward: -20.000,	mean reward: -0.020	[-1.000, 1.000],	mean action: 2.
19349/500000:	episode: 21,	duration: 15.563s,	episode steps: 872,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.024	[-1.000, 0.000],	mean action: 2.
20160/500000:	episode: 22,	duration: 14.346s,	episode steps: 811,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.026	[-1.000, 0.000],	mean action: 2.
21013/500000:	episode: 23,	duration: 14.849s,	episode steps: 853,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action: 2.
21911/500000:	episode: 24,	duration: 15.776s,	episode steps: 898,	steps per second: 57,	episode reward: -20.000,	mean reward: -0.022	[-1.000, 1.000],	mean action: 2.
22735/500000:	episode: 25,	duration: 14.532s,	episode steps: 824,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action: 2.
23638/500000:	episode: 26,	duration: 16.076s,	episode steps: 903,	steps per second: 56,	episode reward: -20.000,	mean reward: -0.022	[-1.000, 1.000],	mean action: 2.
24542/500000:	episode: 27,	duration: 15.969s,	episode steps: 904,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.023	[-1.000, 0.000],	mean action: 2.
25366/500000:	episode: 28,	duration: 14.521s,	episode steps: 824,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action: 2.
26130/500000:	episode: 29,	duration: 13.453s,	episode steps: 764,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action: 2.
26956/500000:	episode: 30,	duration: 14.992s,	episode steps: 826,	steps per second: 55,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action: 2.
27826/500000:	episode: 31,	duration: 15.319s,	episode steps: 870,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.024	[-1.000, 0.000],	mean action: 2.
28770/500000:	episode: 32,	duration: 16.625s,	episode steps: 944,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.022	[-1.000, 0.000],	mean action: 2.
29736/500								

1/8/2020, 5:06 PM

188477/500000:	episode: 218,	duration: 13.425s,	episode steps: 764,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
188477/500000:	episode: 219,	duration: 13.52s,	episode steps: 764,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
189320/500000:	episode: 220,	duration: 14.715s,	episode steps: 843,	steps per second: 57,	episode reward: -20.000,	mean reward: -0.024	[-1.000, 1.000],	mean action:
190084/500000:	episode: 221,	duration: 13.56s,	episode steps: 764,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
190906/500000:	episode: 222,	duration: 14.482s,	episode steps: 822,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.026	[-1.000, 0.000],	mean action:
191670/500000:	episode: 223,	duration: 13.328s,	episode steps: 764,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
192452/500000:	episode: 224,	duration: 14.042s,	episode steps: 782,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
193284/500000:	episode: 225,	duration: 14.724s,	episode steps: 832,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
194242/500000:	episode: 226,	duration: 16.878s,	episode steps: 958,	steps per second: 57,	episode reward: -20.000,	mean reward: -0.021	[-1.000, 1.000],	mean action:
195091/500000:	episode: 227,	duration: 14.740s,	episode steps: 849,	steps per second: 58,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
195883/500000:	episode: 228,	duration: 13.922s,	episode steps: 792,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
196675/500000:	episode: 229,	duration: 13.891s,	episode steps: 792,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
197496/500000:	episode: 230,	duration: 14.497s,	episode steps: 821,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.026	[-1.000, 0.000],	mean action:
198398/500000:	episode: 231,	duration: 16.002s,	episode steps: 902,	steps per second: 56,	episode reward: -20.000,	mean reward: -0.022	[-1.000, 1.000],	mean action:
199238/500000:	episode: 232,	duration: 14.753s,	episode steps: 840,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
200002/500000:	episode: 233,	duration: 13.313s,	episode steps: 764,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
200766/500000:	episode: 234,	duration: 13.414s,	episode steps: 764,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
201530/500000:	episode: 235,	duration: 13.128s,	episode steps: 764,	steps per second: 58,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
202313/500000:	episode: 236,	duration: 13.839s,	episode steps: 783,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
203125/500000:	episode: 237,	duration: 14.639s,	episode steps: 812,	steps per second: 55,	episode reward: -21.000,	mean reward: -0.026	[-1.000, 0.000],	mean action:
203889/500000:	episode: 238,	duration: 13.495s,	episode steps: 764,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
204824/500000:	episode: 239,	duration: 16.482s,	episode steps: 935,	steps per second: 57,	episode reward: -20.000,	mean reward: -0.021	[-1.000, 1.000],	mean action:
205667/500000:	episode: 240,	duration: 14.816s,	episode steps: 843,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
206431/500000:	episode: 241,	duration: 13.511s,	episode steps: 764,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
207273/500000:	episode: 242,	duration: 14.725s,	episode steps: 842,	steps per second: 57,	episode reward: -20.000,	mean reward: -0.024	[-1.000, 1.000],	mean action:
208037/500000:	episode: 243,	duration: 13.440s,	episode steps: 764,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
208849/500000:	episode: 244,	duration: 14.277s,	episode steps: 812,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.026	[-1.000, 0.000],	mean action:
209670/500000:	episode: 245,	duration: 14.761s,	episode steps: 821,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.026	[-1.000, 0.000],	mean action:
210434/500000:	episode: 246,	duration: 13.491s,	episode steps: 764,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
211306/500000:	episode: 247,	duration: 15.271s,	episode steps: 872,	steps per second: 57,	episode reward: -20.000,	mean reward: -0.023	[-1.000, 1.000],	mean action:
212070/500000:	episode: 248,	duration: 13.543s,	episode steps: 764,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
212881/500000:	episode: 249,	duration: 14.569s,	episode steps: 811,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.026	[-1.000, 0.000],	mean action:
213703/500000:	episode: 250,							

374262/500000:	episode: 436,	duration: 13.889s,	episode steps: 792,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
375054/500000:	episode: 437,	duration: 13.797s,	episode steps: 792,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
375838/500000:	episode: 438,	duration: 13.584s,	episode steps: 784,	steps per second: 58,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
376664/500000:	episode: 439,	duration: 14.539s,	episode steps: 826,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
377488/500000:	episode: 440,	duration: 14.579s,	episode steps: 824,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
378341/500000:	episode: 441,	duration: 15.144s,	episode steps: 853,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
379184/500000:	episode: 442,	duration: 15.064s,	episode steps: 843,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
379976/500000:	episode: 443,	duration: 14.072s,	episode steps: 792,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
380768/500000:	episode: 444,	duration: 13.934s,	episode steps: 792,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
381594/500000:	episode: 445,	duration: 14.689s,	episode steps: 826,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
382386/500000:	episode: 446,	duration: 14.125s,	episode steps: 792,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
383179/500000:	episode: 447,	duration: 13.849s,	episode steps: 793,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.026	[-1.000, 0.000],	mean action:
383962/500000:	episode: 448,	duration: 13.738s,	episode steps: 783,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
384788/500000:	episode: 449,	duration: 14.551s,	episode steps: 826,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
385613/500000:	episode: 450,	duration: 14.591s,	episode steps: 825,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
386438/500000:	episode: 451,	duration: 14.406s,	episode steps: 825,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
387262/500000:	episode: 452,	duration: 14.497s,	episode steps: 824,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
388087/500000:	episode: 453,	duration: 14.553s,	episode steps: 825,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
388913/500000:	episode: 454,	duration: 14.582s,	episode steps: 826,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
389705/500000:	episode: 455,	duration: 13.854s,	episode steps: 792,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
390531/500000:	episode: 456,	duration: 14.508s,	episode steps: 826,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
391429/500000:	episode: 457,	duration: 16.009s,	episode steps: 898,	steps per second: 56,	episode reward: -20.000,	mean reward: -0.022	[-1.000, 1.000],	mean action:
392273/500000:	episode: 458,	duration: 14.902s,	episode steps: 844,	steps per second: 57,	episode reward: -20.000,	mean reward: -0.024	[-1.000, 1.000],	mean action:
393122/500000:	episode: 459,	duration: 14.905s,	episode steps: 849,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
393914/500000:	episode: 460,	duration: 14.077s,	episode steps: 792,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
394739/500000:	episode: 461,	duration: 14.540s,	episode steps: 825,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
395582/500000:	episode: 462,	duration: 15.027s,	episode steps: 843,	steps per second: 56,	episode reward: -20.000,	mean reward: -0.024	[-1.000, 1.000],	mean action:
396407/500000:	episode: 463,	duration: 14.833s,	episode steps: 825,	steps per second: 56,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
397199/500000:	episode: 464,	duration: 13.926s,	episode steps: 792,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.027	[-1.000, 0.000],	mean action:
398025/500000:	episode: 465,	duration: 14.551s,	episode steps: 826,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
398851/500000:	episode: 466,	duration: 14.586s,	episode steps: 826,	steps per second: 57,	episode reward: -21.000,	mean reward: -0.025	[-1.000, 0.000],	mean action:
399693/500000:	episode: 467,	duration: 14.975s,	episode steps: 842,	steps per second: 56,	episode reward: -20.000,	mean reward: -0.024	[-1.000, 1.000],	mean action:
400519/500000:	episode: 468,							

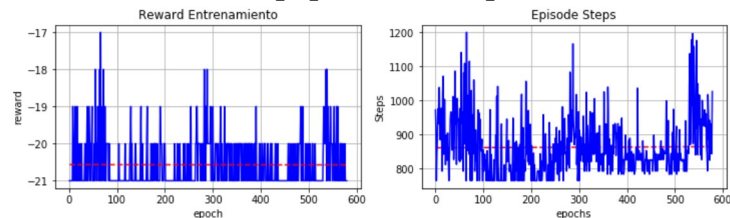
```

467748/500000: episode: 545, duration: 17.963s, episode steps: 1020, steps per second: 57, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action:
468924/500000: episode: 546, duration: 20.822s, episode steps: 1176, steps per second: 56, episode reward: -20.000, mean reward: -0.017 [-1.000, 1.000], mean action:
469841/500000: episode: 547, duration: 16.256s, episode steps: 917, steps per second: 56, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action:
470891/500000: episode: 548, duration: 18.303s, episode steps: 1050, steps per second: 57, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action:
471731/500000: episode: 549, duration: 14.721s, episode steps: 840, steps per second: 57, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
472720/500000: episode: 550, duration: 17.412s, episode steps: 989, steps per second: 57, episode reward: -21.000, mean reward: -0.021 [-1.000, 0.000], mean action:
473651/500000: episode: 551, duration: 16.429s, episode steps: 931, steps per second: 57, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action:
474667/500000: episode: 552, duration: 17.896s, episode steps: 1016, steps per second: 57, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action:
475613/500000: episode: 553, duration: 16.793s, episode steps: 946, steps per second: 56, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action:
476480/500000: episode: 554, duration: 15.236s, episode steps: 867, steps per second: 57, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
477319/500000: episode: 555, duration: 14.839s, episode steps: 839, steps per second: 57, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
478264/500000: episode: 556, duration: 16.708s, episode steps: 945, steps per second: 57, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action:
479090/500000: episode: 557, duration: 14.589s, episode steps: 826, steps per second: 57, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
480090/500000: episode: 558, duration: 17.471s, episode steps: 1000, steps per second: 57, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action:
481035/500000: episode: 559, duration: 16.652s, episode steps: 945, steps per second: 57, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action:
481985/500000: episode: 560, duration: 16.670s, episode steps: 950, steps per second: 57, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action:
482910/500000: episode: 561, duration: 16.294s, episode steps: 925, steps per second: 57, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action:
483839/500000: episode: 562, duration: 16.698s, episode steps: 929, steps per second: 56, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action:
484811/500000: episode: 563, duration: 16.961s, episode steps: 972, steps per second: 57, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action:
485763/500000: episode: 564, duration: 16.721s, episode steps: 952, steps per second: 57, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action:
486686/500000: episode: 565, duration: 16.408s, episode steps: 923, steps per second: 56, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action:
487647/500000: episode: 566, duration: 17.253s, episode steps: 961, steps per second: 56, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action:
488515/500000: episode: 567, duration: 15.392s, episode steps: 868, steps per second: 56, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
489359/500000: episode: 568, duration: 15.007s, episode steps: 844, steps per second: 56, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action:
490152/500000: episode: 569, duration: 14.305s, episode steps: 793, steps per second: 55, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action:
491169/500000: episode: 570, duration: 18.141s, episode steps: 1017, steps per second: 56, episode reward: -21.000, mean reward: -0.021 [-1.000, 0.000], mean action:
492133/500000: episode: 571, duration: 16.830s, episode steps: 964, steps per second: 57, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action:
493019/500000: episode: 572, duration: 15.775s, episode steps: 886, steps per second: 56, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
493961/500000: episode: 573, duration: 16.621s, episode steps: 942, steps per second: 57, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action:
494785/500000: episode: 574, duration: 14.613s, episode steps: 824, steps per second: 56, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
495717/500000: episode: 575, duration: 16.412s, episode steps: 932, steps per second: 57, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action:
496597/500000: episode: 576, duration: 15.490s, episode steps: 880, steps per second: 57, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
497464/500000: episode: 577, duration: 15.288s, episode steps: 867, steps per second: 57, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
498307/500000: episode: 578, duration: 14.853s, episode steps: 843, steps per second: 57, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action:
499333/500000: episode: 579, duration: 18.211s, episode steps: 1026, steps per second: 56, episode reward: -21.000, mean reward: -0.020 [-1.000, 0.000], mean action:

```

done, took 8693.123 seconds

Pesos guardados en: 202001081340_dqn_PongDeterministic-v0_weights.h5f



Testing for 10 episodes ...

```

Episode 1: reward: -19.000, steps: 976
Episode 2: reward: -19.000, steps: 1000
Episode 3: reward: -21.000, steps: 1036
Episode 4: reward: -20.000, steps: 1255
Episode 5: reward: -20.000, steps: 908
Episode 6: reward: -20.000, steps: 946
Episode 7: reward: -21.000, steps: 826
Episode 8: reward: -20.000, steps: 974
Episode 9: reward: -21.000, steps: 876
Episode 10: reward: -20.000, steps: 898
mean_reward: -20.1
mean_steps: 969.5
(~20.1, 969.5)

```