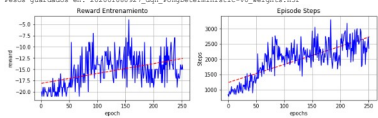


405968/500000: episode: 214, duration: 18.427s, episode steps: 2067, steps per second: 112, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean action: 2.415 [0.000, 5.000], mean observation: 105.555 [87.000, 236.000], loss: 0.002485, mean_absolute_error: 0.603813, mean_q: -0.622493, mean_mpe: 0.100000
405973/500000: episode: 215, duration: 22.708s, episode steps: 2545, steps per second: 112, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean action: 2.496 [0.000, 5.000], mean observation: 105.560 [87.000, 236.000], loss: 0.002458, mean_absolute_error: 0.612189, mean_q: -0.637174, mean_mpe: 0.100000
410971/500000: episode: 216, duration: 22.064s, episode steps: 2458, steps per second: 112, episode reward: -18.000, mean reward: -0.007 [-1.000, 1.000], mean action: 2.362 [0.000, 5.000], mean observation: 105.580 [87.000, 236.000], loss: 0.003013, mean_absolute_error: 0.607482, mean_q: -0.629433, mean_mpe: 0.100000
413296/500000: episode: 217, duration: 20.134s, episode steps: 2323, steps per second: 112, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean action: 2.448 [0.000, 5.000], mean observation: 105.539 [87.000, 236.000], loss: 0.002893, mean_absolute_error: 0.600919, mean_q: -0.624526, mean_mpe: 0.100000
415113/500000: episode: 218, duration: 25.164s, episode steps: 2835, steps per second: 112, episode reward: -15.000, mean reward: -0.009 [-1.000, 1.000], mean action: 2.483 [0.000, 5.000], mean observation: 105.564 [87.000, 236.000], loss: 0.002289, mean_absolute_error: 0.600119, mean_q: -0.616779, mean_mpe: 0.100000
418943/500000: episode: 219, duration: 25.768s, episode steps: 2835, steps per second: 110, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean action: 2.435 [0.000, 5.000], mean observation: 105.561 [87.000, 236.000], loss: 0.002481, mean_absolute_error: 0.614889, mean_q: -0.640528, mean_mpe: 0.100000
420844/500000: episode: 220, duration: 16.868s, episode steps: 1361, steps per second: 112, episode reward: -18.000, mean reward: -0.010 [-1.000, 1.000], mean action: 2.291 [0.000, 5.000], mean observation: 105.518 [87.000, 236.000], loss: 0.003062, mean_absolute_error: 0.614771, mean_q: -0.638188, mean_mpe: 0.100000
423190/500000: episode: 221, duration: 21.078s, episode steps: 2352, steps per second: 112, episode reward: -12.000, mean reward: -0.005 [-1.000, 1.000], mean action: 2.450 [0.000, 5.000], mean observation: 105.545 [87.000, 236.000], loss: 0.002939, mean_absolute_error: 0.609664, mean_q: -0.631847, mean_mpe: 0.100000
425457/500000: episode: 222, duration: 20.483s, episode steps: 2271, steps per second: 111, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean action: 2.354 [0.000, 5.000], mean observation: 105.536 [87.000, 236.000], loss: 0.002417, mean_absolute_error: 0.608014, mean_q: -0.624762, mean_mpe: 0.100000
428327/500000: episode: 223, duration: 25.427s, episode steps: 2805, steps per second: 112, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean action: 2.519 [0.000, 5.000], mean observation: 105.581 [87.000, 236.000], loss: 0.002931, mean_absolute_error: 0.604212, mean_q: -0.619082, mean_mpe: 0.100000
430236/500000: episode: 224, duration: 17.202s, episode steps: 1929, steps per second: 112, episode reward: -18.000, mean reward: -0.009 [-1.000, 1.000], mean action: 2.449 [0.000, 5.000], mean observation: 105.471 [87.000, 236.000], loss: 0.002466, mean_absolute_error: 0.604383, mean_q: -0.619761, mean_mpe: 0.100000
432882/500000: episode: 225, duration: 23.314s, episode steps: 2624, steps per second: 113, episode reward: -10.000, mean reward: -0.004 [-1.000, 1.000], mean action: 2.487 [0.000, 5.000], mean observation: 105.534 [87.000, 236.000], loss: 0.003131, mean_absolute_error: 0.597463, mean_q: -0.609824, mean_mpe: 0.100000
435402/500000: episode: 226, duration: 19.204s, episode steps: 2158, steps per second: 112, episode reward: -13.000, mean reward: -0.008 [-1.000, 1.000], mean action: 2.488 [0.000, 5.000], mean observation: 105.570 [87.000, 236.000], loss: 0.002491, mean_absolute_error: 0.597023, mean_q: -0.612077, mean_mpe: 0.100000
438050/500000: episode: 227, duration: 26.112s, episode steps: 2963, steps per second: 113, episode reward: -12.000, mean reward: -0.004 [-1.000, 1.000], mean action: 2.591 [0.000, 5.000], mean observation: 105.535 [87.000, 236.000], loss: 0.002226, mean_absolute_error: 0.594096, mean_q: -0.609712, mean_mpe: 0.100000
440513/500000: episode: 228, duration: 22.429s, episode steps: 2511, steps per second: 112, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean action: 2.496 [0.000, 5.000], mean observation: 105.501 [87.000, 236.000], loss: 0.002921, mean_absolute_error: 0.597937, mean_q: -0.611775, mean_mpe: 0.100000
442684/500000: episode: 229, duration: 19.313s, episode steps: 2171, steps per second: 112, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean action: 2.400 [0.000, 5.000], mean observation: 105.560 [87.000, 236.000], loss: 0.003109, mean_absolute_error: 0.594752, mean_q: -0.605846, mean_mpe: 0.100000
445571/500000: episode: 230, duration: 25.664s, episode steps: 2887, steps per second: 113, episode reward: -15.000, mean reward: -0.004 [-1.000, 1.000], mean action: 2.598 [0.000, 5.000], mean observation: 105.611 [87.000, 236.000], loss: 0.002920, mean_absolute_error: 0.597121, mean_q: -0.605707, mean_mpe: 0.100000
448457/500000: episode: 231, duration: 26.748s, episode steps: 2996, steps per second: 112, episode reward: -9.000, mean reward: -0.003 [-1.000, 1.000], mean action: 2.487 [0.000, 5.000], mean observation: 105.580 [87.000, 236.000], loss: 0.002586, mean_absolute_error: 0.601094, mean_q: -0.617113, mean_mpe: 0.100000
451126/500000: episode: 232, duration: 22.182s, episode steps: 2541, steps per second: 112, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean action: 2.449 [0.000, 5.000], mean observation: 105.571 [87.000, 236.000], loss: 0.002960, mean_absolute_error: 0.592085, mean_q: -0.605030, mean_mpe: 0.100000
453903/500000: episode: 233, duration: 25.242s, episode steps: 2775, steps per second: 110, episode reward: -10.000, mean reward: -0.005 [-1.000, 1.000], mean action: 2.377 [0.000, 5.000], mean observation: 105.504 [87.000, 236.000], loss: 0.002821, mean_absolute_error: 0.586465, mean_q: -0.597964, mean_mpe: 0.100000
456392/500000: episode: 234, duration: 21.455s, episode steps: 2489, steps per second: 116, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean action: 2.540 [0.000, 5.000], mean observation: 105.508 [87.000, 236.000], loss: 0.002509, mean_absolute_error: 0.591586, mean_q: -0.605706, mean_mpe: 0.100000
458951/500000: episode: 235, duration: 22.743s, episode steps: 2559, steps per second: 115, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean action: 2.478 [0.000, 5.000], mean observation: 105.570 [87.000, 236.000], loss: 0.002486, mean_absolute_error: 0.598939, mean_q: -0.613045, mean_mpe: 0.100000
461544/500000: episode: 236, duration: 22.345s, episode steps: 2593, steps per second: 116, episode reward: -11.000, mean reward: -0.004 [-1.000, 1.000], mean action: 2.555 [0.000, 5.000], mean observation: 105.567 [87.000, 236.000], loss: 0.003418, mean_absolute_error: 0.605939, mean_q: -0.626951, mean_mpe: 0.100000
464030/500000: episode: 237, duration: 20.717s, episode steps: 2366, steps per second: 116, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean action: 2.406 [0.000, 5.000], mean observation: 105.620 [87.000, 236.000], loss: 0.002537, mean_absolute_error: 0.597234, mean_q: -0.609709, mean_mpe: 0.100000
467421/500000: episode: 238, duration: 24.947s, episode steps: 2891, steps per second: 116, episode reward: -8.000, mean reward: -0.003 [-1.000, 1.000], mean action: 2.502 [0.000, 5.000], mean observation: 105.615 [87.000, 236.000], loss: 0.002365, mean_absolute_error: 0.595729, mean_q: -0.604932, mean_mpe: 0.100000
469179/500000: episode: 239, duration: 15.111s, episode steps: 1758, steps per second: 116, episode reward: -14.000, mean reward: -0.009 [-1.000, 1.000], mean action: 2.325 [0.000, 5.000], mean observation: 105.614 [87.000, 236.000], loss: 0.002380, mean_absolute_error: 0.590809, mean_q: -0.598032, mean_mpe: 0.100000
470883/500000: episode: 240, duration: 15.574s, episode steps: 1804, steps per second: 116, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean action: 2.440 [0.000, 5.000], mean observation: 105.553 [87.000, 236.000], loss: 0.003591, mean_absolute_error: 0.580427, mean_q: -0.584607, mean_mpe: 0.100000
473249/500000: episode: 241, duration: 19.747s, episode steps: 2286, steps per second: 116, episode reward: -15.000, mean reward: -0.007 [-1.000, 1.000], mean action: 2.451 [0.000, 5.000], mean observation: 105.586 [87.000, 236.000], loss: 0.003216, mean_absolute_error: 0.580206, mean_q: -0.591854, mean_mpe: 0.100000
475925/500000: episode: 242, duration: 22.760s, episode steps: 2656, steps per second: 117, episode reward: -19.000, mean reward: -0.007 [-1.000, 1.000], mean action: 2.333 [0.000, 5.000], mean observation: 105.560 [87.000, 236.000], loss: 0.002732, mean_absolute_error: 0.594345, mean_q: -0.613193, mean_mpe: 0.100000
478116/500000: episode: 243, duration: 18.821s, episode steps: 2191, steps per second: 116, episode reward: -17.000, mean reward: -0.008 [-1.000, 1.000], mean action: 2.440 [0.000, 5.000], mean observation: 105.614 [87.000, 236.000], loss: 0.002664, mean_absolute_error: 0.608224, mean_q: -0.632300, mean_mpe: 0.100000
480519/500000: episode: 244, duration: 20.776s, episode steps: 2403, steps per second: 116, episode reward: -14.000, mean reward: -0.007 [-1.000, 1.000], mean action: 2.380 [0.000, 5.000], mean observation: 105.581 [87.000, 236.000], loss: 0.002843, mean_absolute_error: 0.614428, mean_q: -0.641746, mean_mpe: 0.100000
483150/500000: episode: 245, duration: 22.078s, episode steps: 2581, steps per second: 117, episode reward: -17.000, mean reward: -0.007 [-1.000, 1.000], mean action: 2.375 [0.000, 5.000], mean observation: 105.554 [87.000, 236.000], loss: 0.003118, mean_absolute_error: 0.618760, mean_q: -0.644635, mean_mpe: 0.100000
484924/500000: episode: 246, duration: 15.814s, episode steps: 1824, steps per second: 116, episode reward: -17.000, mean reward: -0.009 [-1.000, 1.000], mean action: 2.454 [0.000, 5.000], mean observation: 105.617 [87.000, 236.000], loss: 0.002521, mean_absolute_error: 0.623160, mean_q: -0.650545, mean_mpe: 0.100000
487730/500000: episode: 247, duration: 20.386s, episode steps: 2368, steps per second: 116, episode reward: -16.000, mean reward: -0.007 [-1.000, 1.000], mean action: 2.457 [0.000, 5.000], mean observation: 105.511 [87.000, 236.000], loss: 0.002324, mean_absolute_error: 0.621613, mean_q: -0.647653, mean_mpe: 0.100000
489524/500000: episode: 248, duration: 19.496s, episode steps: 2232, steps per second: 114, episode reward: -14.000, mean reward: -0.006 [-1.000, 1.000], mean action: 2.392 [0.000, 5.000], mean observation: 105.545 [87.000, 236.000], loss: 0.002955, mean_absolute_error: 0.621630, mean_q: -0.646834, mean_mpe: 0.100000
491909/500000: episode: 249, duration: 20.518s, episode steps: 2385, steps per second: 116, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean action: 2.517 [0.000, 5.000], mean observation: 105.512 [87.000, 236.000], loss: 0.002889, mean_absolute_error: 0.615214, mean_q: -0.639495, mean_mpe: 0.100000
494440/500000: episode: 250, duration: 22.077s, episode steps: 2531, steps per second: 114, episode reward: -15.000, mean reward: -0.004 [-1.000, 1.000], mean action: 2.527 [0.000, 5.000], mean observation: 105.575 [87.000, 236.000], loss: 0.003103, mean_absolute_error: 0.614845, mean_q: -0.638475, mean_mpe: 0.100000
497147/500000: episode: 251, duration: 23.715s, episode steps: 2727, steps per second: 117, episode reward: -14.000, mean reward: -0.005 [-1.000, 1.000], mean action: 2.420 [0.000, 5.000], mean observation: 105.578 [87.000, 236.000], loss: 0.002952, mean_absolute_error: 0.612791, mean_q: -0.637075, mean_mpe: 0.100000
499521/500000: episode: 252, duration: 20.402s, episode steps: 2314, steps per second: 115, episode reward: -15.000, mean reward: -0.006 [-1.000, 1.000], mean action: 2.442 [0.000, 5.000], mean observation: 105.580 [87.000, 236.000], loss: 0.002451, mean_absolute_error: 0.613736, mean_q: -0.637111, mean_mpe: 0.100000

done, took 4374.001 seconds
Pwaa guardado em: 20201019T07:00_PongDeteministic-v0_weights.455



Twisting for 10 episodes:

- Episode 1: reward: -14.000, steps: 2462
 - Episode 2: reward: -15.000, steps: 2088
 - Episode 3: reward: -12.000, steps: 2904
 - Episode 4: reward: -8.000, steps: 2435
 - Episode 5: reward: -11.000, steps: 2977
 - Episode 6: reward: -9.000, steps: 2909
 - Episode 7: reward: -11.000, steps: 2643
 - Episode 8: reward: -12.000, steps: 2530
 - Episode 9: reward: -16.000, steps: 2048
 - Episode 10: reward: -14.000, steps: 2374
- mean reward: -12.3
mean_mpe: 2564.4
(-12.3, 2564.9)