

[illegible]

9/1/20 17:04

1008769/1500000: episode: 453, duration: 10.466s, episode steps: 1890, steps per second: 181, episode reward: 19.000, mean reward: 0.010 (-1.000, 1.000), mean action: 2.440 (0.000, 5.000), mean observation: 105.568 (87.000, 236.000), loss: 0.001691, mean\_absolute\_error: 0.798622, mean\_q: 1.020880, mean\_episode\_reward: 19.000000

101041/1500000: episode: 454, duration: 10.269s, episode steps: 1872, steps per second: 182, episode reward: 19.000, mean reward: 0.018 (-1.000, 1.000), mean action: 2.241 (0.000, 5.000), mean observation: 105.494 (87.000, 236.000), loss: 0.001820, mean\_absolute\_error: 0.798371, mean\_q: 1.021488, mean\_episode\_reward: 19.000000

101135/1500000: episode: 455, duration: 11.774s, episode steps: 2014, steps per second: 182, episode reward: 12.000, mean reward: 0.005 (-1.000, 1.000), mean action: 2.439 (0.000, 5.000), mean observation: 105.579 (87.000, 236.000), loss: 0.001720, mean\_absolute\_error: 0.804812, mean\_q: 1.031543, mean\_episode\_reward: 12.000000

101548/1500000: episode: 456, duration: 12.419s, episode steps: 2211, steps per second: 181, episode reward: 12.000, mean reward: 0.007 (-1.000, 1.000), mean action: 2.376 (0.000, 5.000), mean observation: 105.548 (87.000, 236.000), loss: 0.001790, mean\_absolute\_error: 0.803582, mean\_q: 1.031339, mean\_episode\_reward: 12.000000

1016176/1500000: episode: 457, duration: 10.353s, episode steps: 2190, steps per second: 181, episode reward: 12.000, mean reward: 0.004 (-1.000, 1.000), mean action: 2.366 (0.000, 5.000), mean observation: 105.568 (87.000, 236.000), loss: 0.001876, mean\_absolute\_error: 0.789761, mean\_q: 1.009214, mean\_episode\_reward: 12.000000

102096/1500000: episode: 458, duration: 10.434s, episode steps: 2198, steps per second: 181, episode reward: 12.000, mean reward: 0.004 (-1.000, 1.000), mean action: 2.440 (0.000, 5.000), mean observation: 105.534 (87.000, 236.000), loss: 0.001800, mean\_absolute\_error: 0.752561, mean\_q: 0.705054, mean\_episode\_reward: 12.000000

102145/1500000: episode: 459, duration: 12.128s, episode steps: 2139, steps per second: 181, episode reward: 12.000, mean reward: 0.007 (-1.000, 1.000), mean action: 2.399 (0.000, 5.000), mean observation: 105.565 (87.000, 236.000), loss: 0.001876, mean\_absolute\_error: 0.735464, mean\_q: 0.742127, mean\_episode\_reward: 12.000000

102422/1500000: episode: 460, duration: 9.727s, episode steps: 1757, steps per second: 181, episode reward: 20.000, mean reward: 0.011 (-1.000, 1.000), mean action: 2.360 (0.000, 5.000), mean observation: 105.558 (87.000, 236.000), loss: 0.001643, mean\_absolute\_error: 0.747189, mean\_q: 0.761669, mean\_episode\_reward: 20.000000

102621/1500000: episode: 461, duration: 11.084s, episode steps: 2199, steps per second: 181, episode reward: 16.000, mean reward: 0.009 (-1.000, 1.000), mean action: 2.458 (0.000, 5.000), mean observation: 105.572 (87.000, 236.000), loss: 0.001634, mean\_absolute\_error: 0.752222, mean\_q: 0.764216, mean\_episode\_reward: 16.000000

102898/1500000: episode: 462, duration: 10.828s, episode steps: 1971, steps per second: 182, episode reward: 17.000, mean reward: 0.009 (-1.000, 1.000), mean action: 2.435 (0.000, 5.000), mean observation: 105.515 (87.000, 236.000), loss: 0.001384, mean\_absolute\_error: 0.774959, mean\_q: 0.795512, mean\_episode\_reward: 17.000000

103109/1500000: episode: 463, duration: 12.115s, episode steps: 2191, steps per second: 181, episode reward: 16.000, mean reward: 0.008 (-1.000, 1.000), mean action: 2.385 (0.000, 5.000), mean observation: 105.464 (87.000, 236.000), loss: 0.001203, mean\_absolute\_error: 0.793164, mean\_q: 0.751262, mean\_episode\_reward: 16.000000

103347/1500000: episode: 464, duration: 12.899s, episode steps: 2371, steps per second: 182, episode reward: 14.000, mean reward: 0.006 (-1.000, 1.000), mean action: 2.429 (0.000, 5.000), mean observation: 105.591 (87.000, 236.000), loss: 0.001360, mean\_absolute\_error: 0.797574, mean\_q: 0.720548, mean\_episode\_reward: 14.000000

103505/1500000: episode: 465, duration: 10.829s, episode steps: 1978, steps per second: 181, episode reward: 19.000, mean reward: 0.018 (-1.000, 1.000), mean action: 2.420 (0.000, 5.000), mean observation: 105.588 (87.000, 236.000), loss: 0.001476, mean\_absolute\_error: 0.782841, mean\_q: 0.707659, mean\_episode\_reward: 19.000000

103776/1500000: episode: 466, duration: 12.994s, episode steps: 2375, steps per second: 182, episode reward: 17.000, mean reward: 0.007 (-1.000, 1.000), mean action: 2.413 (0.000, 5.000), mean observation: 105.531 (87.000, 236.000), loss: 0.001767, mean\_absolute\_error: 0.778948, mean\_q: 0.709929, mean\_episode\_reward: 17.000000

103949/1500000: episode: 467, duration: 11.833s, episode steps: 2149, steps per second: 181, episode reward: 14.000, mean reward: 0.007 (-1.000, 1.000), mean action: 2.411 (0.000, 5.000), mean observation: 105.580 (87.000, 236.000), loss: 0.001760, mean\_absolute\_error: 0.771851, mean\_q: 0.795134, mean\_episode\_reward: 14.000000

104182/1500000: episode: 468, duration: 10.884s, episode steps: 1942, steps per second: 182, episode reward: 17.000, mean reward: 0.009 (-1.000, 1.000), mean action: 2.367 (0.000, 5.000), mean observation: 105.551 (87.000, 236.000), loss: 0.001897, mean\_absolute\_error: 0.780892, mean\_q: 1.006036, mean\_episode\_reward: 17.000000

104423/1500000: episode: 469, duration: 12.076s, episode steps: 2351, steps per second: 182, episode reward: 15.000, mean reward: 0.006 (-1.000, 1.000), mean action: 2.703 (0.000, 5.000), mean observation: 105.568 (87.000, 236.000), loss: 0.001544, mean\_absolute\_error: 0.798439, mean\_q: 0.719526, mean\_episode\_reward: 15.000000

104610/1500000: episode: 470, duration: 10.314s, episode steps: 1872, steps per second: 182, episode reward: 15.000, mean reward: 0.010 (-1.000, 1.000), mean action: 2.314 (0.000, 5.000), mean observation: 105.565 (87.000, 236.000), loss: 0.001638, mean\_absolute\_error: 0.787905, mean\_q: 1.009062, mean\_episode\_reward: 15.000000

104823/1500000: episode: 471, duration: 11.809s, episode steps: 2128, steps per second: 182, episode reward: 14.000, mean reward: 0.007 (-1.000, 1.000), mean action: 2.421 (0.000, 5.000), mean observation: 105.538 (87.000, 236.000), loss: 0.001410, mean\_absolute\_error: 0.805214, mean\_q: 1.028590, mean\_episode\_reward: 14.000000

1050138/1500000: episode: 472, duration: 10.373s, episode steps: 1945, steps per second: 185, episode reward: 16.000, mean reward: 0.009 (-1.000, 1.000), mean action: 2.367 (0.000, 5.000), mean observation: 105.504 (87.000, 236.000), loss: 0.001220, mean\_absolute\_error: 0.813075, mean\_q: 1.041218, mean\_episode\_reward: 16.000000

105193/1500000: episode: 473, duration: 9.739s, episode steps: 1757, steps per second: 180, episode reward: 20.000, mean reward: 0.011 (-1.000, 1.000), mean action: 2.427 (0.000, 5.000), mean observation: 105.433 (87.000, 236.000), loss: 0.001446, mean\_absolute\_error: 0.834144, mean\_q: 1.003143, mean\_episode\_reward: 20.000000

105371/1500000: episode: 474, duration: 11.709s, episode steps: 2142, steps per second: 182, episode reward: 16.000, mean reward: 0.008 (-1.000, 1.000), mean action: 2.375 (0.000, 5.000), mean observation: 105.492 (87.000, 236.000), loss: 0.001109, mean\_absolute\_error: 0.846424, mean\_q: 1.077150, mean\_episode\_reward: 16.000000

105579/1500000: episode: 475, duration: 11.259s, episode steps: 2046, steps per second: 182, episode reward: 16.000, mean reward: 0.008 (-1.000, 1.000), mean action: 2.352 (0.000, 5.000), mean observation: 105.511 (87.000, 236.000), loss: 0.001027, mean\_absolute\_error: 0.869550, mean\_q: 1.105797, mean\_episode\_reward: 16.000000

105731/1500000: episode: 476, duration: 11.709s, episode steps: 2142, steps per second: 182, episode reward: 16.000, mean reward: 0.007 (-1.000, 1.000), mean action: 2.446 (0.000, 5.000), mean observation: 105.568 (87.000, 236.000), loss: 0.001140, mean\_absolute\_error: 0.867532, mean\_q: 1.107181, mean\_episode\_reward: 16.000000

106024/1500000: episode: 477, duration: 12.696s, episode steps: 2329, steps per second: 182, episode reward: 12.000, mean reward: 0.006 (-1.000, 1.000), mean action: 2.329 (0.000, 5.000), mean observation: 105.574 (87.000, 236.000), loss: 0.001303, mean\_absolute\_error: 0.866185, mean\_q: 1.095908, mean\_episode\_reward: 12.000000

106205/1500000: episode: 478, duration: 10.211s, episode steps: 1854, steps per second: 182, episode reward: 19.000, mean reward: 0.010 (-1.000, 1.000), mean action: 2.474 (0.000, 5.000), mean observation: 105.588 (87.000, 236.000), loss: 0.001503, mean\_absolute\_error: 0.851499, mean\_q: 1.174312, mean\_episode\_reward: 19.000000

106428/1500000: episode: 479, duration: 11.763s, episode steps: 2112, steps per second: 181, episode reward: 16.000, mean reward: 0.008 (-1.000, 1.000), mean action: 2.382 (0.000, 5.000), mean observation: 105.588 (87.000, 236.000), loss: 0.001540, mean\_absolute\_error: 0.841467, mean\_q: 1.003854, mean\_episode\_reward: 16.000000

106604/1500000: episode: 480, duration: 10.392s, episode steps: 1856, steps per second: 182, episode reward: 19.000, mean reward: 0.010 (-1.000, 1.000), mean action: 2.341 (0.000, 5.000), mean observation: 105.481 (87.000, 236.000), loss: 0.001296, mean\_absolute\_error: 0.847712, mean\_q: 1.170486, mean\_episode\_reward: 19.000000

106742/1500000: episode: 481, duration: 9.715s, episode steps: 1750, steps per second: 181, episode reward: 20.000, mean reward: 0.011 (-1.000, 1.000), mean action: 2.522 (0.000, 5.000), mean observation: 105.555 (87.000, 236.000), loss: 0.001392, mean\_absolute\_error: 0.859026, mean\_q: 1.092541, mean\_episode\_reward: 20.000000

107049/1500000: episode: 482, duration: 14.289s, episode steps: 2607, steps per second: 181, episode reward: 16.000, mean reward: 0.006 (-1.000, 1.000), mean action: 2.507 (0.000, 5.000), mean observation: 105.584 (87.000, 236.000), loss: 0.001252, mean\_absolute\_error: 0.848078, mean\_q: 1.174098, mean\_episode\_reward: 16.000000

107122/1500000: episode: 483, duration: 14.724s, episode steps: 2682, steps per second: 182, episode reward: 11.000, mean reward: 0.005 (-1.000, 1.000), mean action: 2.447 (0.000, 5.000), mean observation: 105.529 (87.000, 236.000), loss: 0.001694, mean\_absolute\_error: 0.821020, mean\_q: 1.064412, mean\_episode\_reward: 11.000000

107438/1500000: episode: 484, duration: 9.859s, episode steps: 1786, steps per second: 181, episode reward: 20.000, mean reward: 0.011 (-1.000, 1.000), mean action: 2.312 (0.000, 5.000), mean observation: 105.430 (87.000, 236.000), loss: 0.001434, mean\_absolute\_error: 0.817276, mean\_q: 1.041252, mean\_episode\_reward: 20.000000

107693/1500000: episode: 485, duration: 11.438s, episode steps: 2075, steps per second: 182, episode reward: 16.000, mean reward: 0.009 (-1.000, 1.000), mean action: 2.414 (0.000, 5.000), mean observation: 105.500 (87.000, 236.000), loss: 0.001308, mean\_absolute\_error: 0.820358, mean\_q: 1.040543, mean\_episode\_reward: 16.000000

107927/1500000: episode: 486, duration: 12.699s, episode steps: 2378, steps per second: 182, episode reward: 16.000, mean reward: 0.007 (-1.000, 1.000), mean action: 2.421 (0.000, 5.000), mean observation: 105.535 (87.000, 236.000), loss: 0.001398, mean\_absolute\_error: 0.819278, mean\_q: 1.049776, mean\_episode\_reward: 16.000000

108118/1500000: episode: 487, duration: 11.427s, episode steps: 2047, steps per second: 179, episode reward: 17.000, mean reward: 0.008 (-1.000, 1.000), mean action: 2.344 (0.000, 5.000), mean observation: 105.538 (87.000, 236.000), loss: 0.001476, mean\_absolute\_error: 0.842425, mean\_q: 1.076045, mean\_episode\_reward: 17.000000

108407/1500000: episode: 488, duration: 10.385s, episode steps: 2198, steps per second: 182, episode reward: 14.000, mean reward: 0.005 (-1.000, 1.000), mean action: 2.568 (0.000, 5.000), mean observation: 105.564 (87.000, 236.000), loss: 0.001570, mean\_absolute\_error: 0.852161, mean\_q: 1.098238, mean\_episode\_reward: 14.000000

108598/1500000: episode: 489, duration: 10.726s, episode steps: 1882, steps per second: 182, episode reward: 19.000, mean reward: 0.018 (-1.000, 1.000), mean action: 2.319 (0.000, 5.000), mean observation: 105.528 (87.000, 236.000), loss: 0.001686, mean\_absolute\_error: 0.838957, mean\_q: 1.056483, mean\_episode\_reward: 19.000000

108739/1500000: episode: 490, duration: 10.801s, episode steps: 1841, steps per second: 181, episode reward: 21.000, mean reward: 0.011 (0.000, 1.000), mean action: 2.409 (0.000, 5.000), mean observation: 105.614 (87.000, 236.000), loss: 0.001349, mean\_absolute\_error: 0.841190, mean\_q: 1.009553, mean\_episode\_reward: 21.000000

109003/1500000: episode: 491, duration: 12.709s, episode steps: 2384, steps per second: 181, episode reward: 12.000, mean reward: 0.007 (-1.000, 1.000), mean action: 2.285 (0.000, 5.000), mean observation: 105.584 (87.000, 236.000), loss: 0.001442, mean\_absolute\_error: 0.847702, mean\_q: 1.174745, mean\_episode\_reward: 12.000000

109158/1500000: episode: 492, duration: 10.824s, episode steps: 1897, steps per second: 182, episode reward: 12.000, mean reward: 0.011 (-1.000, 1.000), mean action: 2.313 (0.000, 5.000), mean observation: 105.564 (87.000, 236.000), loss: 0.001481, mean\_absolute\_error: 0.849466, mean\_q: 1.178619, mean\_episode\_reward: 12.000000

109498/1500000: episode: 493, duration: 14.554s, episode steps: 2611, steps per second: 180, episode reward: 12.000, mean reward: 0.005 (-1.000, 1.000), mean action: 2.429 (0.000, 5.000), mean observation: 105.499 (87.000, 236.000), loss: 0.001450, mean\_absolute\_error: 0.832785, mean\_q: 1.066221, mean\_episode\_reward: 12.000000

109642/1500000: episode: 494, duration: 12.348s, episode steps: 2232, steps per second: 182, episode reward: 16.000, mean reward: 0.008 (-1.000, 1.000), mean action: 2.385 (0.000, 5.000), mean observation: 105.617 (87.000, 236.000), loss: 0.001178, mean\_absolute\_error: 0.838969, mean\_q: 1.077455, mean\_episode\_reward: 16.000000

109882/1500000: episode: 495, duration: 11.210s, episode steps: 2038, steps per second: 181, episode reward: 17.000, mean reward: 0.008 (-1.000, 1.000), mean action: 2.461 (0.000, 5.000), mean observation: 105.551 (87.000, 236.000), loss: 0.002210, mean\_absolute\_error: 0.813866, mean\_q: 1.042483, mean\_episode\_reward: 17.000000

110102/1500000: episode: 496, duration: 12.816s, episode steps: 2138, steps per second: 179, episode reward: 12.000, mean reward: 0.009 (-1.000, 1.000), mean action: 2.391 (0.000, 5.000), mean observation: 105.586 (87.000, 236.000), loss: 0.001280, mean\_absolute\_error: 0.811091, mean\_q: 1.040014, mean\_episode\_reward: 12.000000

110314/1500000: episode: 497, duration: 14.767s, episode steps: 2692, steps per second: 182, episode reward: 14.000, mean reward: 0.005 (-1.000, 1.000), mean action: 2.608 (0.000, 5.000), mean observation: 105.531 (87.000, 236.000), loss: 0.001621, mean\_absolute\_error: 0.799192, mean\_q: 1.024126, mean\_episode\_reward: 14.000000

110610/1500000: episode: 498, duration: 16.565s, episode steps: 2611, steps per second: 179, episode reward: 12.000, mean reward: 0.005 (-1.000, 1.000), mean action: 2.455 (0.000, 5.000), mean observation: 105.504 (87.000, 236.000), loss: 0.001681, mean\_absolute\_error: 0.780597, mean\_q: 1.010266, mean\_episode\_reward: 12.000000

110914/1500000: episode: 499, duration: 15.765s, episode steps: 2819, steps per second: 179, episode reward: -1.000, mean reward: -0.000 (-1.000, 1.000), mean action: 2.527 (0.000, 5.000), mean observation: 105.580 (87.000, 236.000), loss: 0.001732, mean\_absolute\_error: 0.783349, mean\_q: 1.005833, mean\_episode\_reward: -1.000000

111297/1500000: episode: 500, duration: 12.416s, episode steps: 2243, steps per second: 179, episode reward: 12.000, mean reward: 0.008 (-1.000, 1.000), mean action: 2.484 (0.000, 5.000), mean observation: 105.513 (87.000, 236.000), loss: 0.001787, mean\_absolute\_error: 0.743387, mean\_q: 0.848813, mean\_episode\_reward: 12.000000

111374/1500000: episode: 501, duration: 12.728s, episode steps: 2307, steps per second: 181, episode reward: 17.000, mean reward: 0.007 (-1.000, 1.000), mean action: 2.748 (0.000, 5.000), mean observation: 105.513 (87.000, 236.000), loss: 0.001771, mean\_absolute\_error: 0.754002, mean\_q: 0.962322, mean\_episode\_reward: 17.000000

111571/1500000: episode: 502, duration: 10.259s, episode steps: 1828, steps per second: 182, episode reward: 12.000, mean reward: 0.018 (-1.000, 1.000), mean action: 2.438 (0.000, 5.000), mean observation: 105.527 (87.000, 236.000), loss: 0.001485, mean\_absolute\_error: 0.779485, mean\_q: 0.795218, mean\_episode\_reward: 12.000000

111751/1500000: episode: 503, duration: 11.107s, episode steps: 2039, steps per second: 181, episode reward: 15.000, mean reward: 0.009 (-1.000, 1.000), mean action: 2.313 (0.000, 5.000), mean observation: 105.543 (87.000, 236.000), loss: 0.001217, mean\_absolute\_error: 0.827710, mean\_q: 1.021841, mean\_episode\_reward: 15.000000

111942/1500000: episode: 504, duration: 11.536s, episode steps: 2103, steps per second: 182, episode reward: 12.000, mean reward: 0.009 (-1.000, 1.000), mean action: 2.378 (0.000, 5.000), mean observation: 105.531 (87.000, 236.000), loss: 0.001182, mean\_absolute\_error: 0.818575, mean\_q: 1.044710, mean\_episode\_reward: 12.000000

112178/1500000: episode: 505, duration: 11.706s, episode steps: 2136, steps per second: 182, episode reward: 20.000, mean reward: 0.009 (-1.000, 1.000), mean action: 2.432 (0.000, 5.000), mean observation: 105.492 (87.000, 236.000), loss: 0.001238, mean\_absolute\_error: 0.845366, mean\_q: 1.008819, mean\_episode\_reward: 20.000000

112383/1500000: episode: 506, duration: 11.252s, episode steps: 2037, steps per second: 181, episode reward: 20.000, mean reward: 0.010 (-1.000, 1.000), mean action: 2.368 (0.000, 5.000), mean observation: 105.470 (87.000, 236.000), loss: 0.001165, mean\_absolute\_error: 0.855681, mean\_q: 1.051212, mean\_episode\_reward: 20.000000

done, took 4231.251 seconds

Penna guardaré em: 20201112\_02\_dpg\_PumpDeterministic-v0\_weights.h5

Training for 10 episodes

Episode 1: reward: 20.000, steps: 1759

Episode 2: reward: 21.000, steps: 1661

Episode 3: reward: 21.000, steps: 1742

Episode 4: reward: 18.000, steps: 2018

Episode 5: reward: 18.000, steps: 1995

Episode 6: reward: 18.000, steps: 1874

Episode 7: reward: 18.000, steps: 2047

Episode 8: reward: 20.000, steps: 1802

Episode 9: reward: 18.000, steps: 2003

Episode 10: reward: 18.000, steps: 2283

mean\_reward: 19.0

mean\_steps: 1918.4