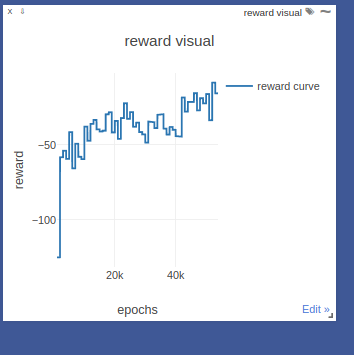
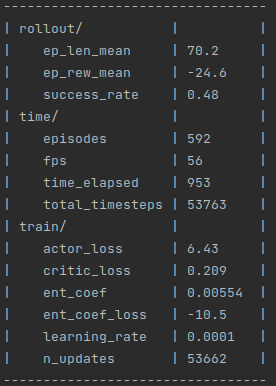
14/12/2022

model = SAC(MultiInputPolicy, learning\_rate=1e-4, gamma=0.95, env=env, verbose=1)

the above parameter is the same for all training here.

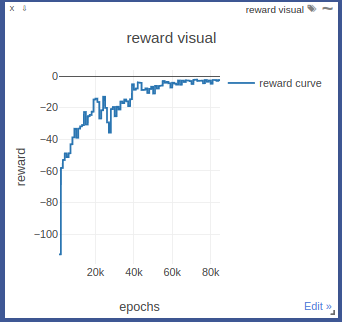
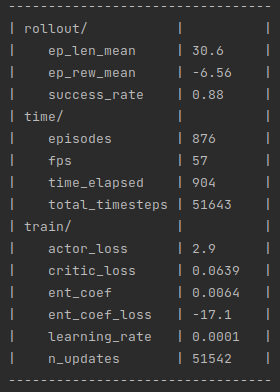
trying to train the robot with the simplest seeting, reaching without regulation. The action is set to +-pi max, but limit to 0.3 sec every step. Success rate is about 0.5 after 50000 epochs. The training processed 3956 steps in 60 seconds.



First we’ll try turn off the rendering and see if that increases the processing speed.

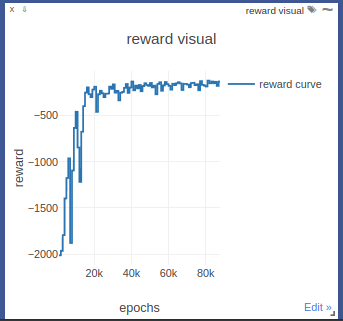
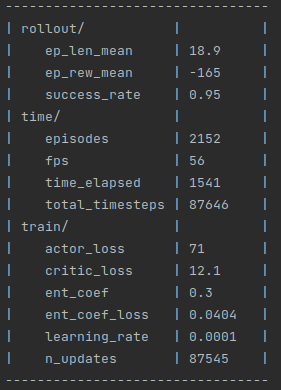
2719 steps processed in 60sec after turning off the render, not improving the training speed.

After changing the action limit to 0.1 sec per step, the training perfomace improved significantly. All other parameters are same with the previous one.



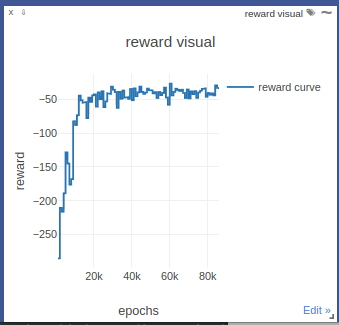
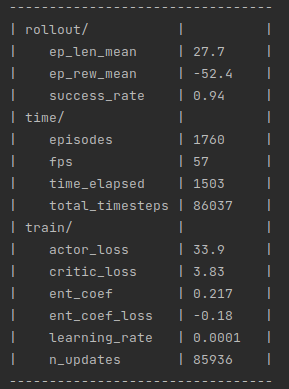
Reach + joint regulation:

Using the ? Function for distance reward, joint regulation coefficient -1, distance coefficient -200, no collision detection.



next we will try to change the distance coefficient to -20.

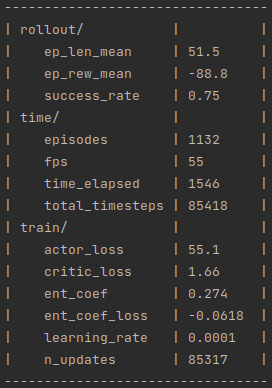
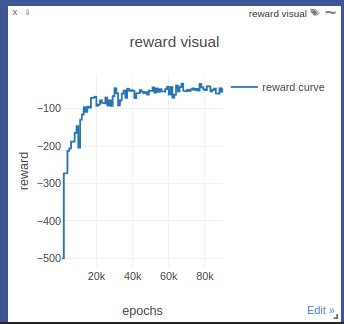
Can’t spot a big difference between this trial and the previous one



Now we try to change the regulation weight on each joint to encourage less motion on the big joint

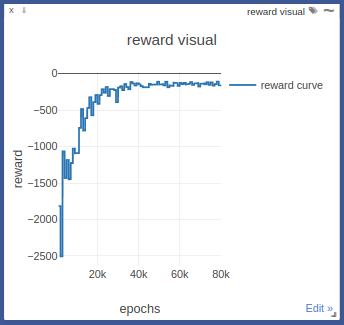
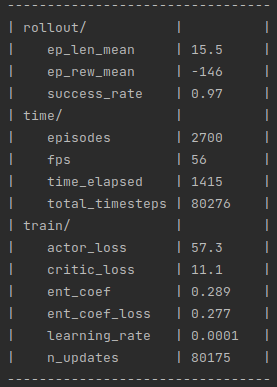
np.sum(np.dot(np.square(self.robot.get\_action()), [2.0, 1.8, 1.6, 1.4, 1.2, 1.0])) \* self.action\_weight

the peformance is not improved by doing this



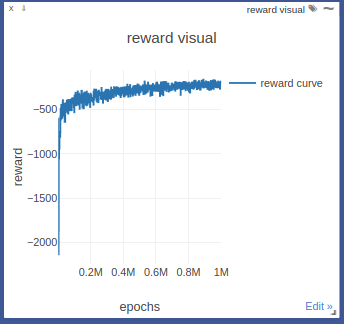
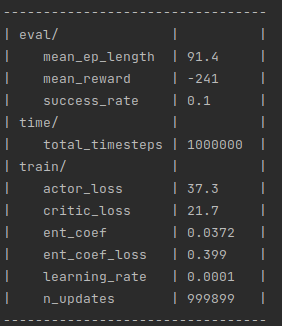
Now we try to add collision detection to the reach + regulation training, once the robot collide with table or track, -200 reward will be applied.

Performance improved after adding collision detection



15/12/2022

Try to train the robot with orientation matching, using the collision weight -200, translational distance weight -200, orientational distance weight -4. The model seems can still improve after 1 million steps. However, the success rate is only 10%. I suspect there are orientations cannot be achieved.



wrote a program that controls the robotic arm using inverse kinematic only, after 100 exxperiments and 100 trials per experiment, we get a successful rate 14.25%. This kind of proves that reach 10% success rate with reinformcement learning is almost hitting the plato and the reward curve tend to flat. However, a 6DoF robot should be able to any point in the 3D space, we need to confirm the success rate with something else, for example, traditional algorithms like RRT connect.



Do another training with the translational distance weight -160, orientational distance weight -4. (the optimal ratio we carried out from 10 experiments)