

## Figures

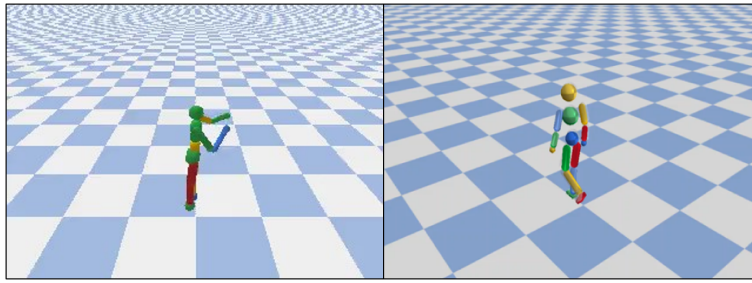


Figure 1: Two versions of the PyBullet Humanoid: HumanoidBulletEnv-v0 (left) and PyBulletDeepMimicEnv (right).

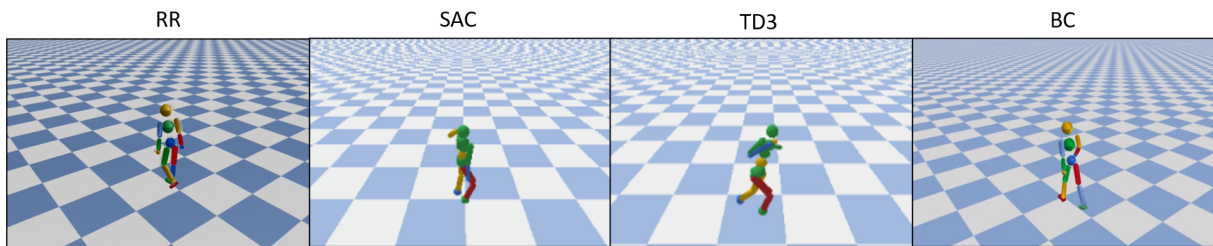


Figure 2: Each model developed a unique gait during the learning process.

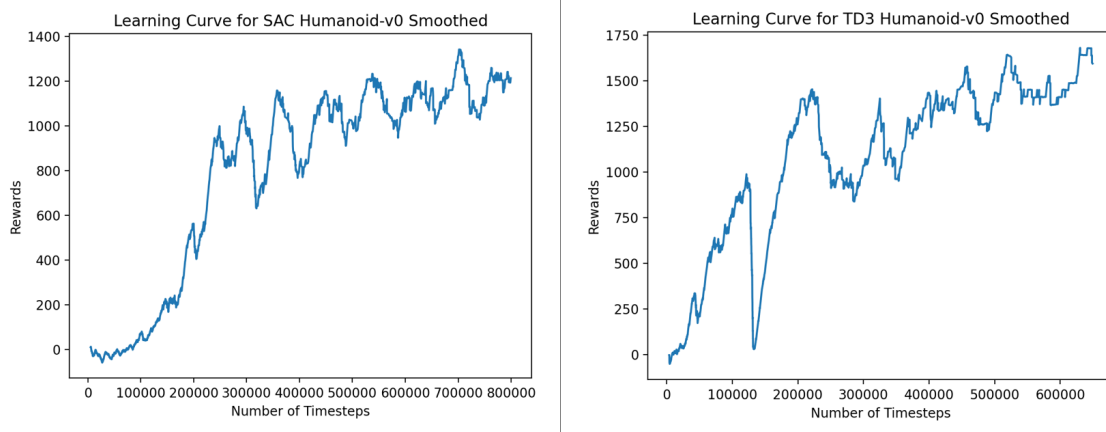


Figure 3: Rewards earned in a sample training session of SAC and TD3

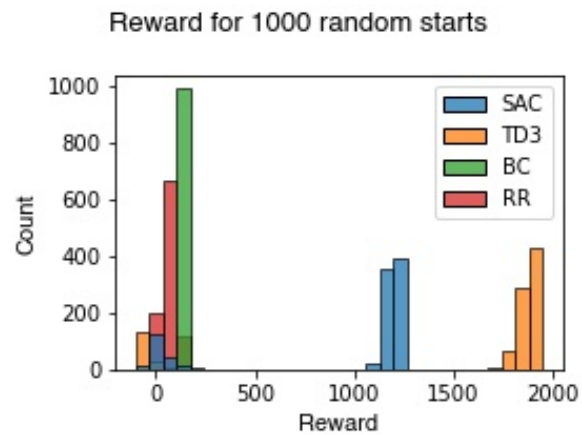


Figure 4: Distribution of the total reward per episode.

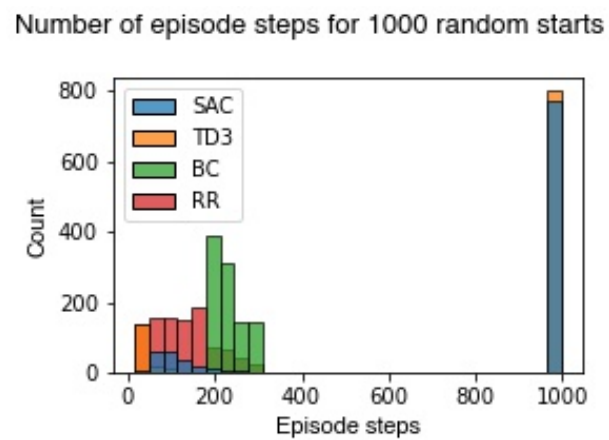


Figure 5: Distribution of the total number of episode steps.