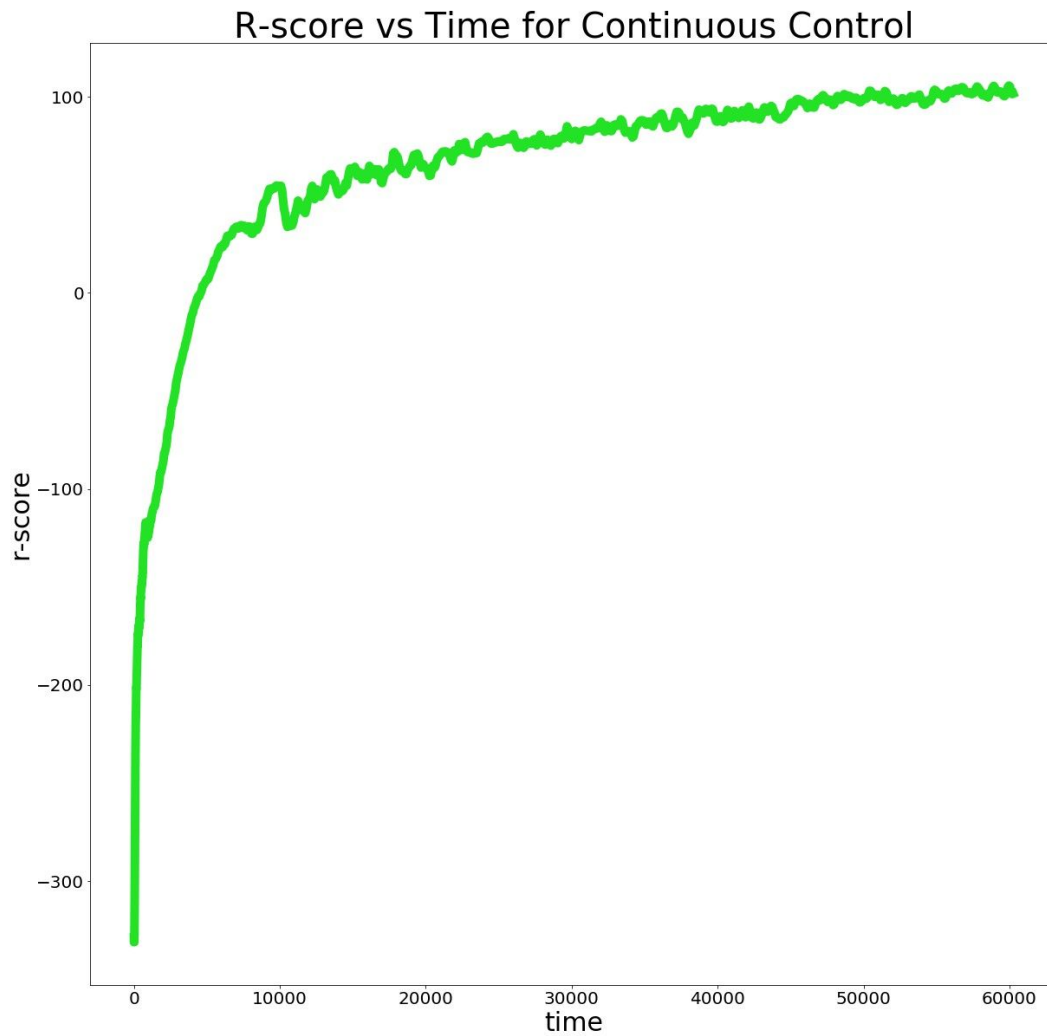


HW7 Reinforcement Learning Report for Continuous Control:

We'll use the same GA3C code, but with minor modifications to make it work on a continuous control task - Lunar Lander Continuous.

Submit the learning curve, i.e. plot of average return (RScore from GA3C) vs. time.

The plots of Average Return (RScore from GA3C) vs. time for the Lunar Lander Continuous is shown below:



The policy is able to make good forward progress, starting at the initial time steps. The learning curve is fairly steep at the start and then becomes more and more flat as time progresses after passing the R-Score threshold of 0.

The curve converges to around an R-Score of 100 (approximately) as more and more timesteps are taken.