Effect of initialisation on the "greedy policy" reward from experimental state random experimental 

timesteps

Rewards