Change process of episode reward for PPO for instances of size n=10