CartPole-v1 - Training

replay_capacity: 10000
gamma: 0.99
lr: 0.07
batch: 256
update_every: 5
init_temp: 5.0
zero_temp_at: 800

Temperature

Score (Raw)
Score (Avg. 10)