Training Progress by Hyperparameters - Rewards per Episode γ : 0.9, ϵ -decay: 0.95 0 --250 Seward (smoothed) -750 - -1000 - -1250 -500 -1500 -1750 LR: 0.002 200 1000 400 600 800 Episodes γ: 0.95, ε-decay: 0.95 0 --200 Reward (smoothed) -400 -600 -800 -1000-1200 — LR: 0.002 200 400 600 800 1000 Episodes γ: 0.99, ε-decay: 0.95 0 -**-250** Reward (smoothed) -750 - -750 - -1000 **-**1250 · -1500 · — LR: 0.002 200 1000 400 600 800 Episodes