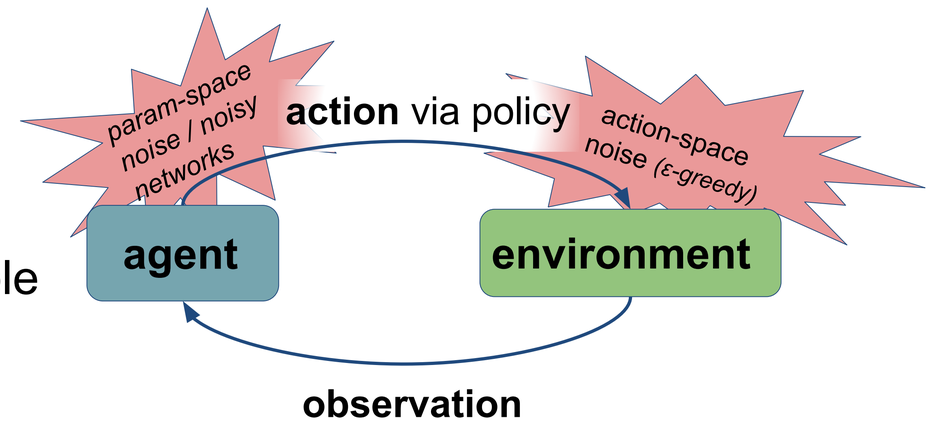
Network Noise for Exploration

Deep Learning Lab Course 2017/18: Badhreesh M Rao, David-Elias Künstle

Project goal

- Extend OpenAI baselines by including Noisy Networks
- Compare training with other exploration strategies using Deep Double Q-Learning Networks (DDQN)



Exploration

- Exploration is needed for the agent to move to new states in its environment
- This aids in exploiting the environment by taking the best actions in every state, to reach its goal
- Three different strategies considered

1) ε-greedy

- Classic exploration strategy (simple, established)
- Deterministic action choice which maximizes expected reward (Q)
- With probability of *ε*: choose random action instead
- Decrease *ε* over time

2) Parameter Space Noise

- Noise added to weights and biases of action-value-model (MLP)
- Adapt noise to keep a distance *δ* between perturbed and non-perturbed policy
- Reduce *δ* over time

3) Noisy Networks

- Noise added to weights and biases of action-value-model (MLP)
- Parameters (μ, σ) of noise learned with gradient descent along with network weights and biases

Experiments

- Run on Mountain Car environment (Figure 1)
- Only *ε-greedy* solved environment perfectly after long training (approx. 350,000 steps)
- *Parameter Space Noise* learned slower and performed worse (Figure 2)
- *Noisy Networks* learned slower than *ε-greedy*
- *Noisy Networks* reached similar reward than *ε-greedy* for short training
- Noisy Networks don't need a manually defined *ε-* or *δ*-decay function

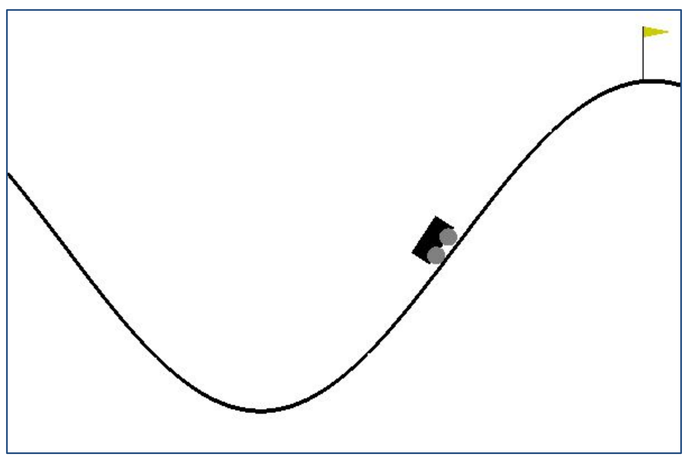


Figure 1: The underpowered car has to reach the flag as fast as possible.

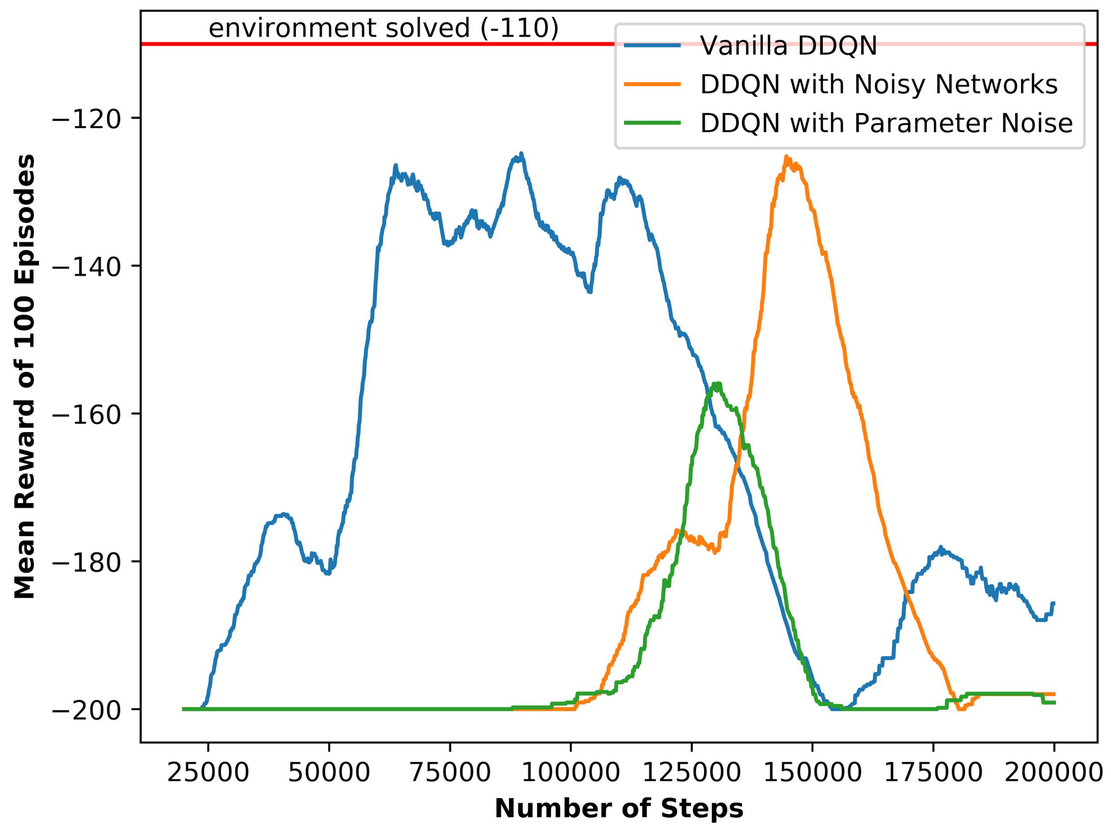


Figure 2: Reward while training DDQN reinforcement learning model in Mountain Car environment with different exploration strategies.

References: Fortunato, M. *et al.* Noisy Networks for Exploration. (2017). Plappert, M. *et al.* Parameter Space Noise for Exploration. (2017). Dhariwal, P. *et al. OpenAI Baselines*. (GitHub, 2017);
Code & Data: https://github.com/dekuenstle/dl_lab_2017/tree/master/final_project_robotics