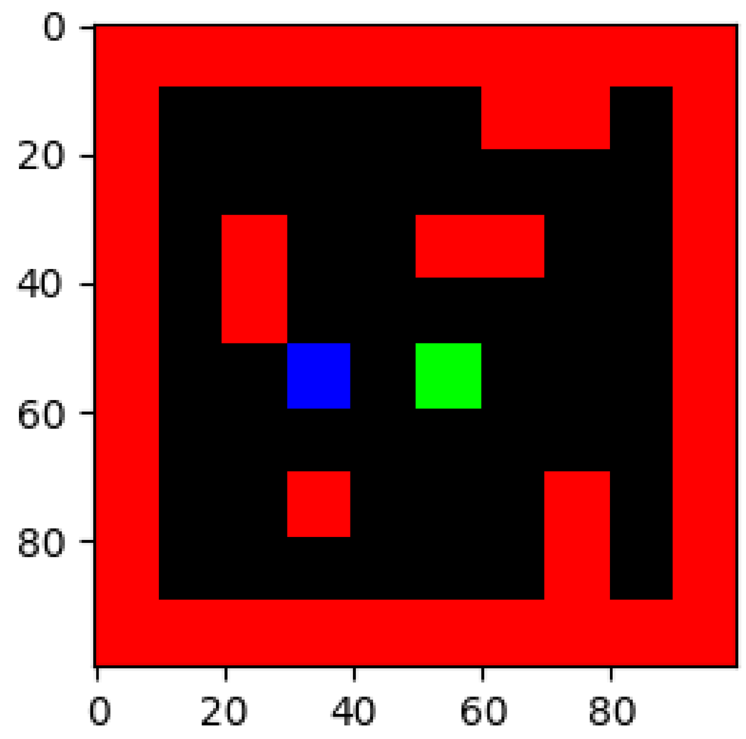
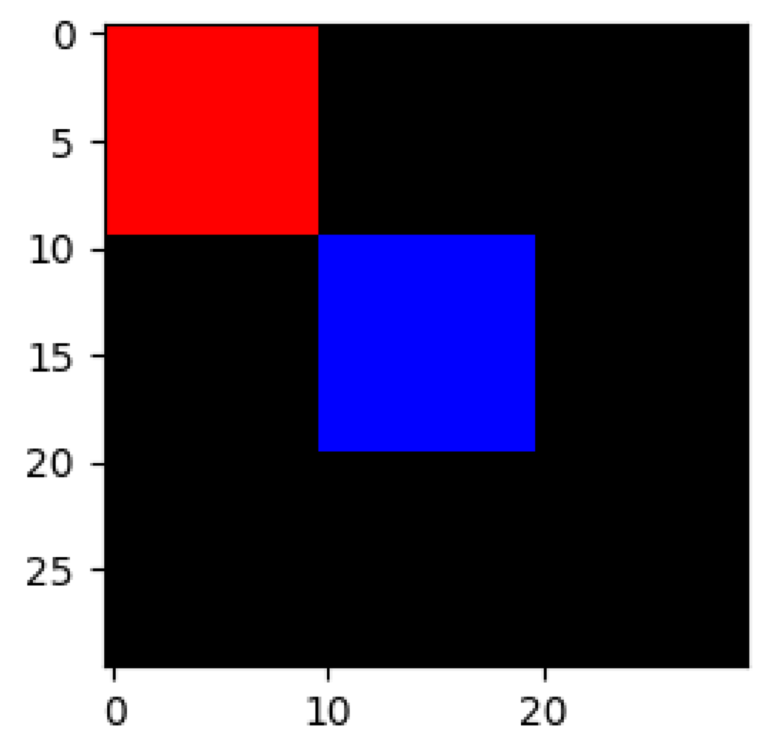
Parameter space noise exploration with target nets in DQN

Fabien Jenne, Johannes Engler

Project Outline

- Visual Planning using Deep-Q-Learning
- Stage 1:
  Training- & target net to stabilize learning
  ε-greedy exploration
- Stage 2:
  Noisy nets for parameter space exploration





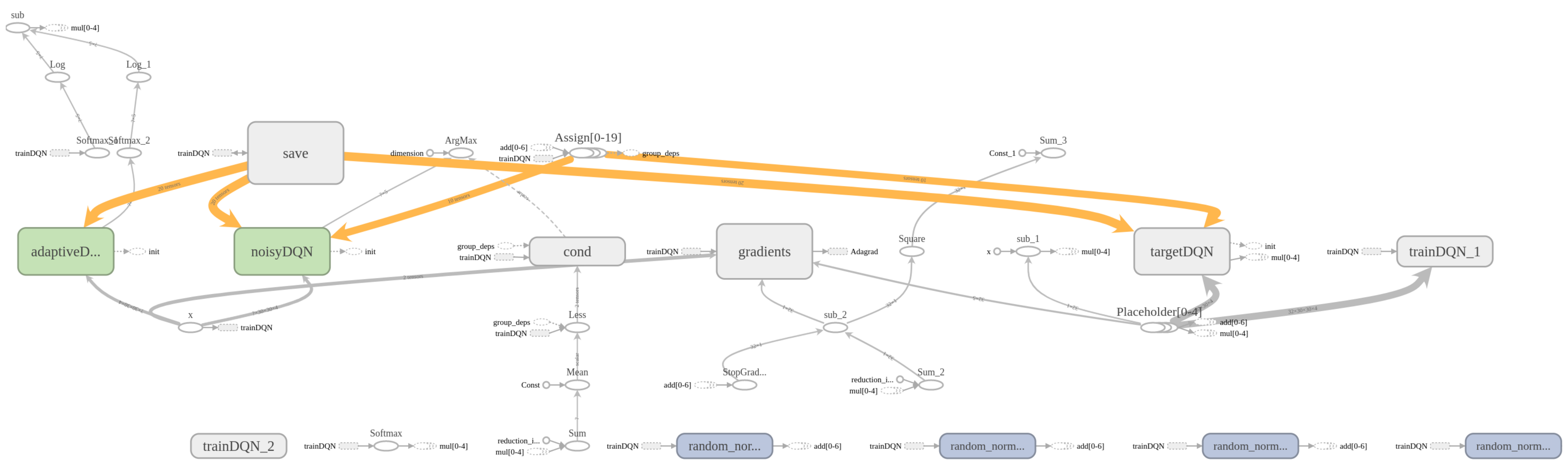
Baseline

1) Convolution: 5x5x32, stride 2, ReLu
2) Convolution: 5x5x32, stride 2, ReLu
3) FullyConnected: 128 nodes, ReLu
4) Dropout: 30%
5) FullyConnected: 64 nodes, ReLu
6) Dropout: 30%
7) FullyConnected: 5 nodes, no activation

- Adagrad Optimizer, learning rate = $5 * 10^{-4}$
- $\gamma = 0.8$
- $\varepsilon = [0.9, 0.1]$

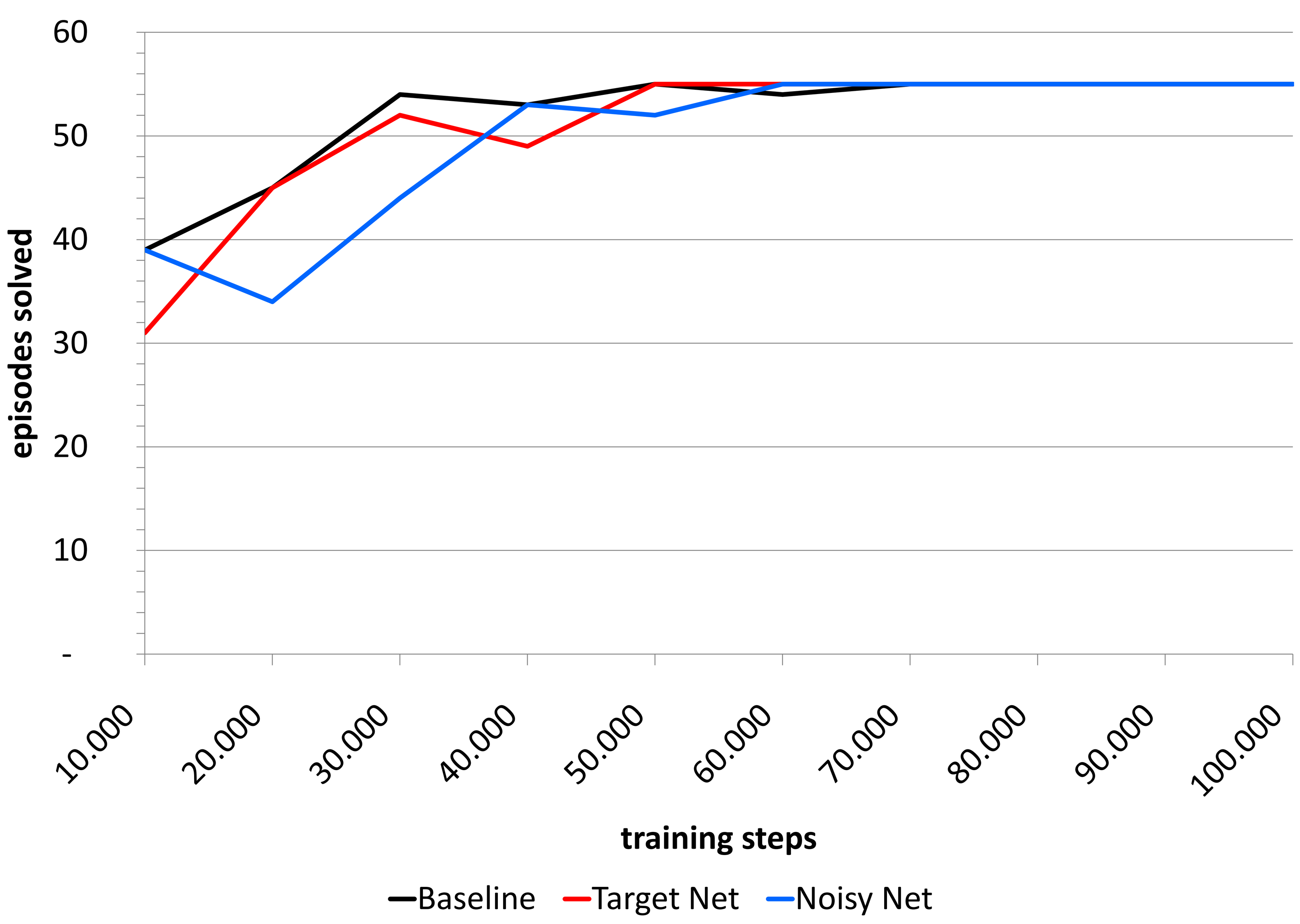
Target Net

- splitting into training & target network
- target network predicts Q_{next}
- weight update every τ steps
- stabilized learning
- slightly better performance



Noisy Nets

- noise added to weights of fully-connected layers
- predicted action is state-dependent
- more thorough exploration compared to ε-greedy
- adaptive noise scaling



Performance

- Target Net slightly better than Baseline
- Noisy Net yields similar results with different approach, slightly slower
- high CPU load
- trained on: i7 8700, 16 GB RAM, GTX 1080