Miniproject

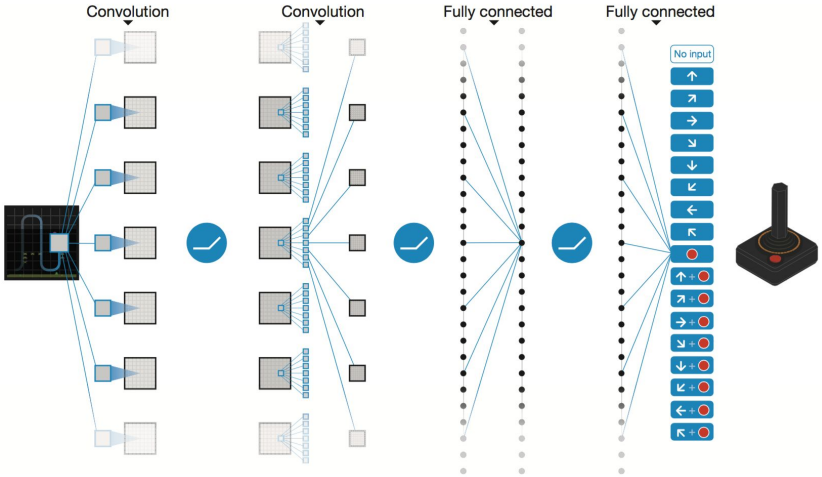
Franz Scherr
Nikolaus Kopp
Thomas Bohnstingl

Content

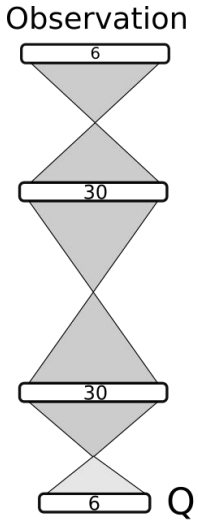
- Algorithm
- Environments
- Results
- Problems

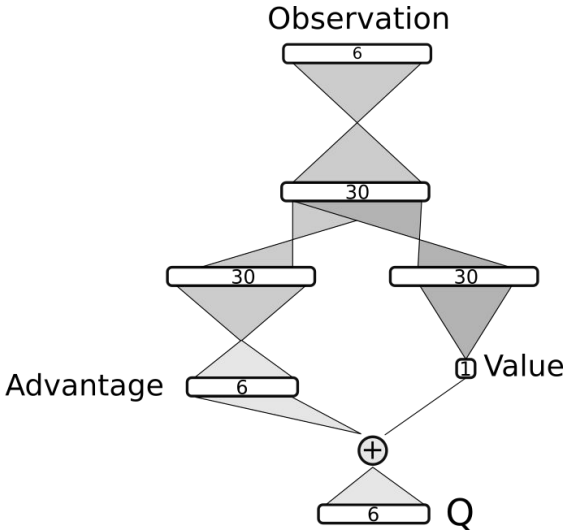
Algorithm

- Deep Q-learning with experience replay
  - Convolutional layers and fully connected layers
- Double Q-learning



Algorithm cont





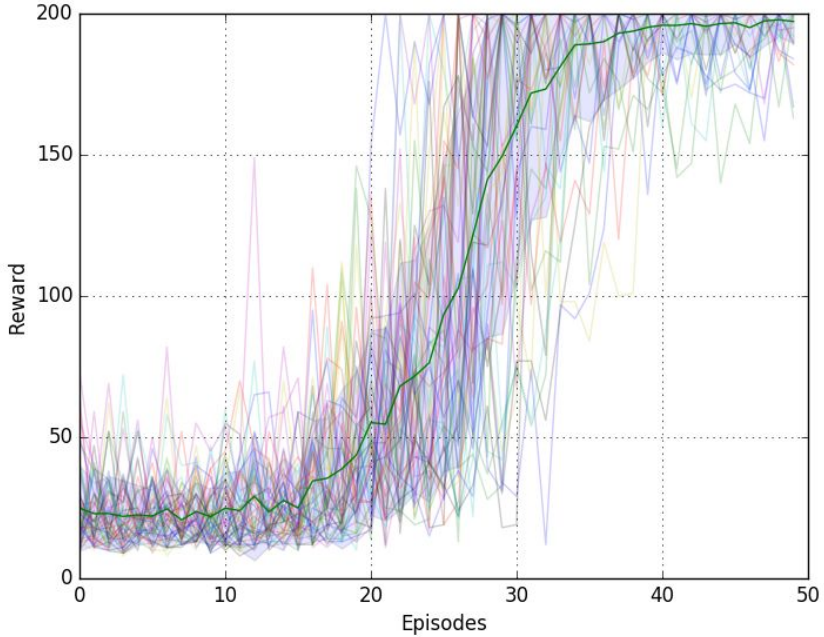
- Duelling Q-learning
  - Split into value and action part
  - Better generalization
  - Basis to implement model-based learning
- Improved Action Selection
  - Epsilon greedy can be improved
  - Use Boltzmann distribution and softmax over the Q-values
  - Temperature decay over time
  - Problem varying reward

Environments

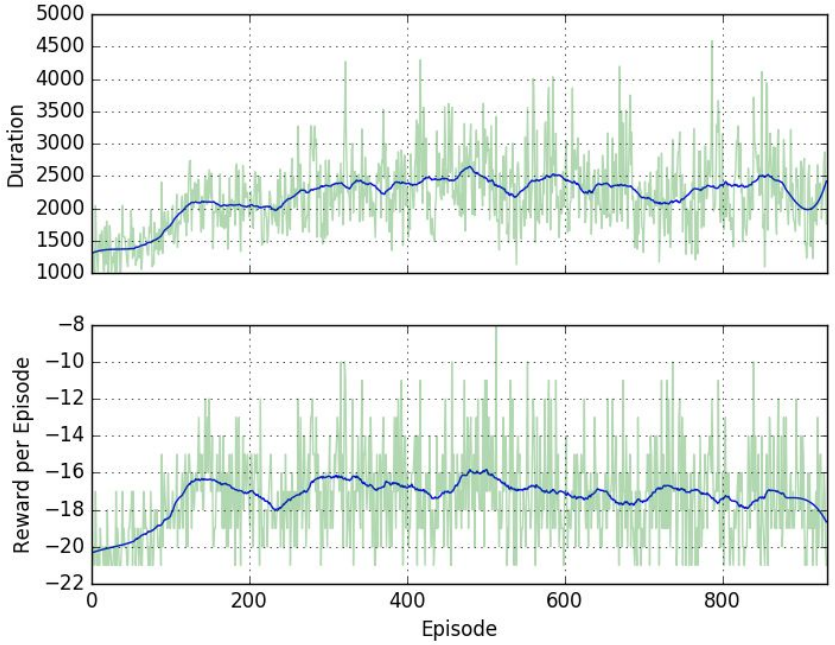
- DQN without convolutional layer
  - Cart Pole
  - Pong
- Convolutional layer
  - MNist data set
  - Pong

Results

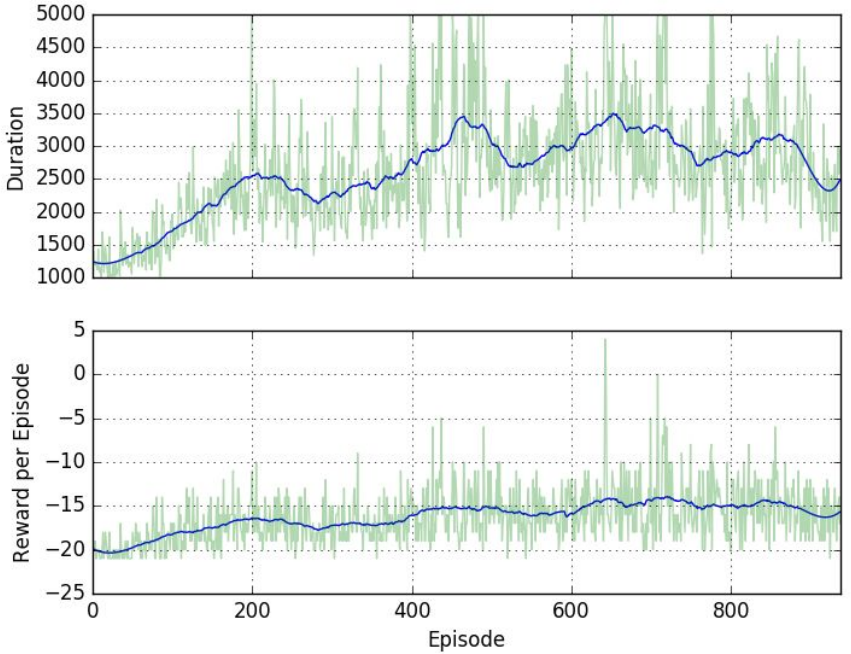
- Cartpole solved in < 50 iterations episodes
- Pong score of -15
- MNist score of 99.2%
- Conv. supervised training with 6 features yielded no results



Results cont



without Duelling Q-learning



with Duelling Q-learning

Problems

- Parameter finding
- Limited resources for training
- Problem with feature extraction in pong