

Hyperparameter search for playing StarCraft II with DQN with non-spatial features

and performance comparison between DQN, DDQN and Ape-X

Florian Soulier
23. August 2019

Agenda

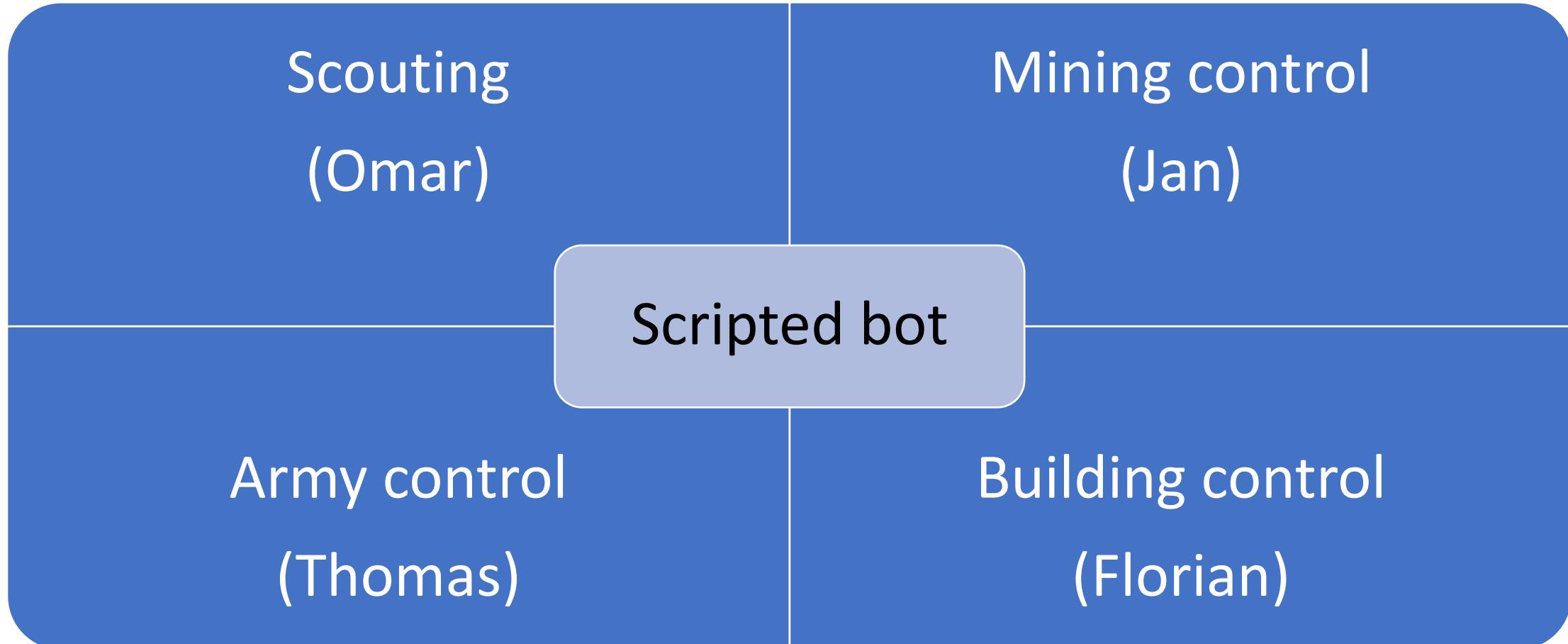
- Motivation
- StarCraft II
- Scripted bot
- The learning environment
- DQN Theory
- Results

Motivation

- Dabbling in machine learning topics in “Scientific Computing” lecture
- Recent success in playing Atari games, AlphaGo and DotA 2
- Publication of „StarCraft II: A new challenge for Reinforcement Learning“
- StarCraft II is a cool game ;-)



Implementation of a scripted bot



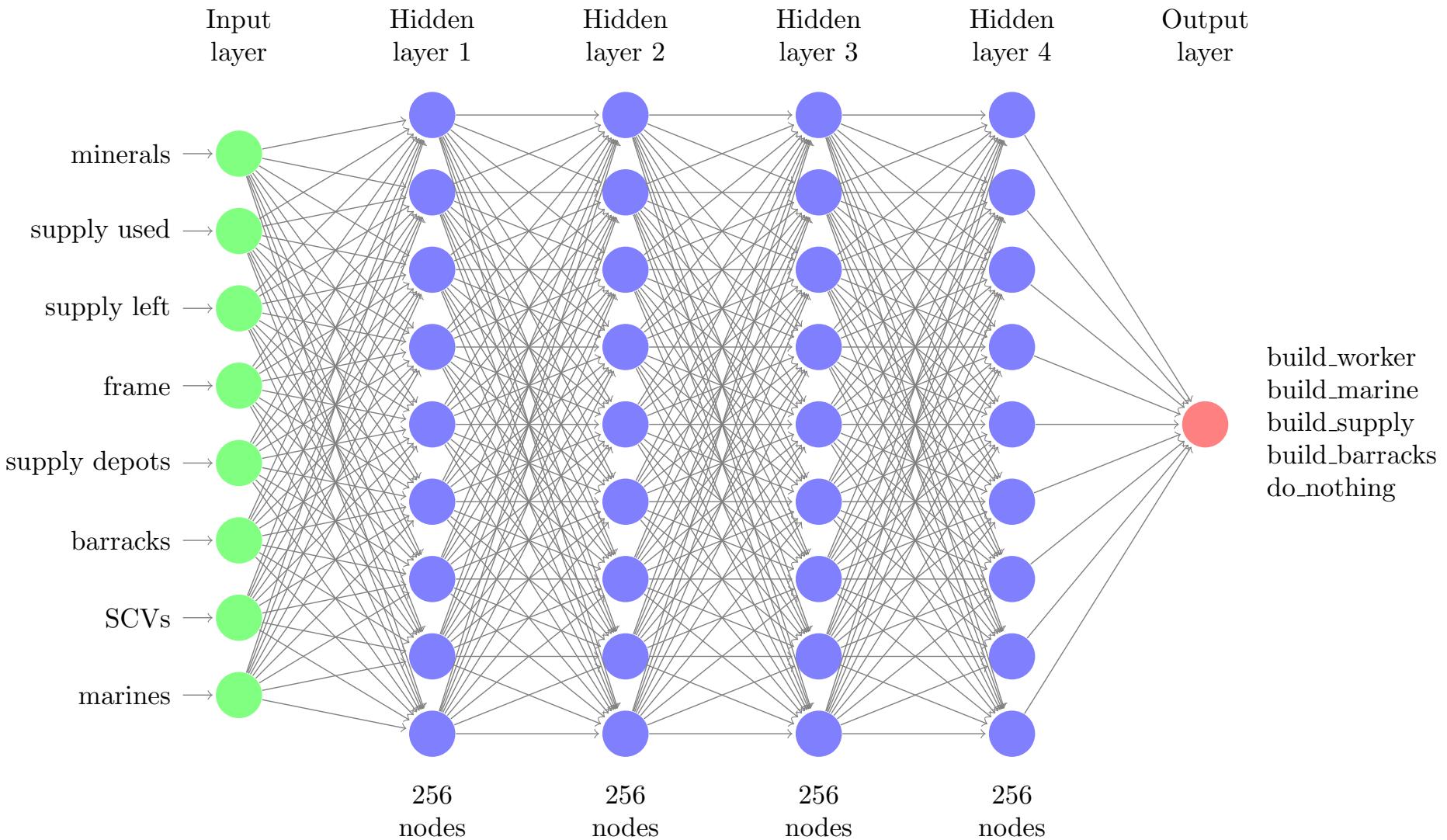
Conclusions from this group project

- Complete game too big to solve -> make problem easier to solve
 - **Omit spatial features** to spare computation power
 - **Use Mini-games** to reduce complexity
 - Use RLLib for **training parallelization** to spare computation time
 - Use **atomic actions** to reduce complexity

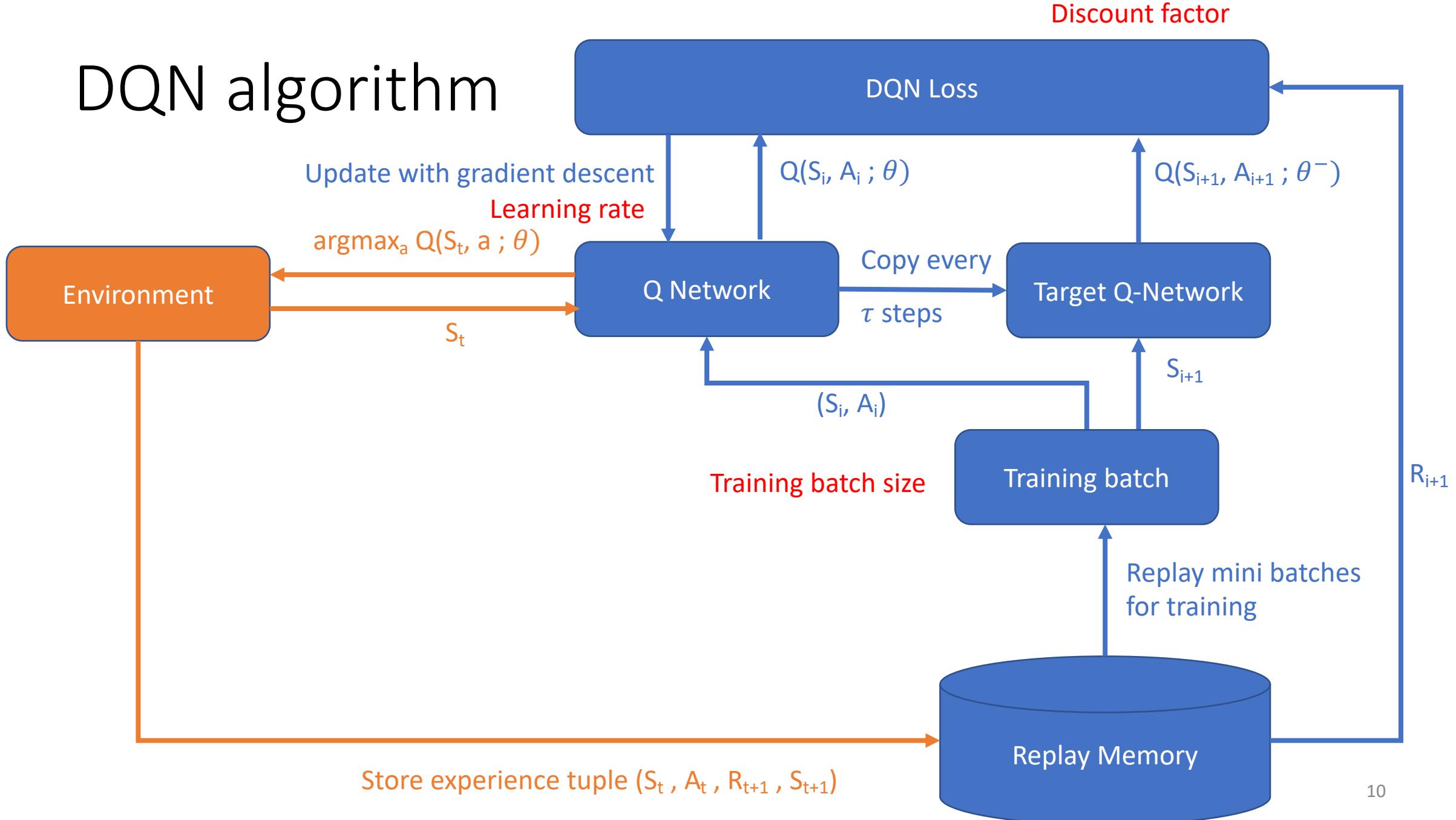
The scenario “BuildMarines”



The learning environment

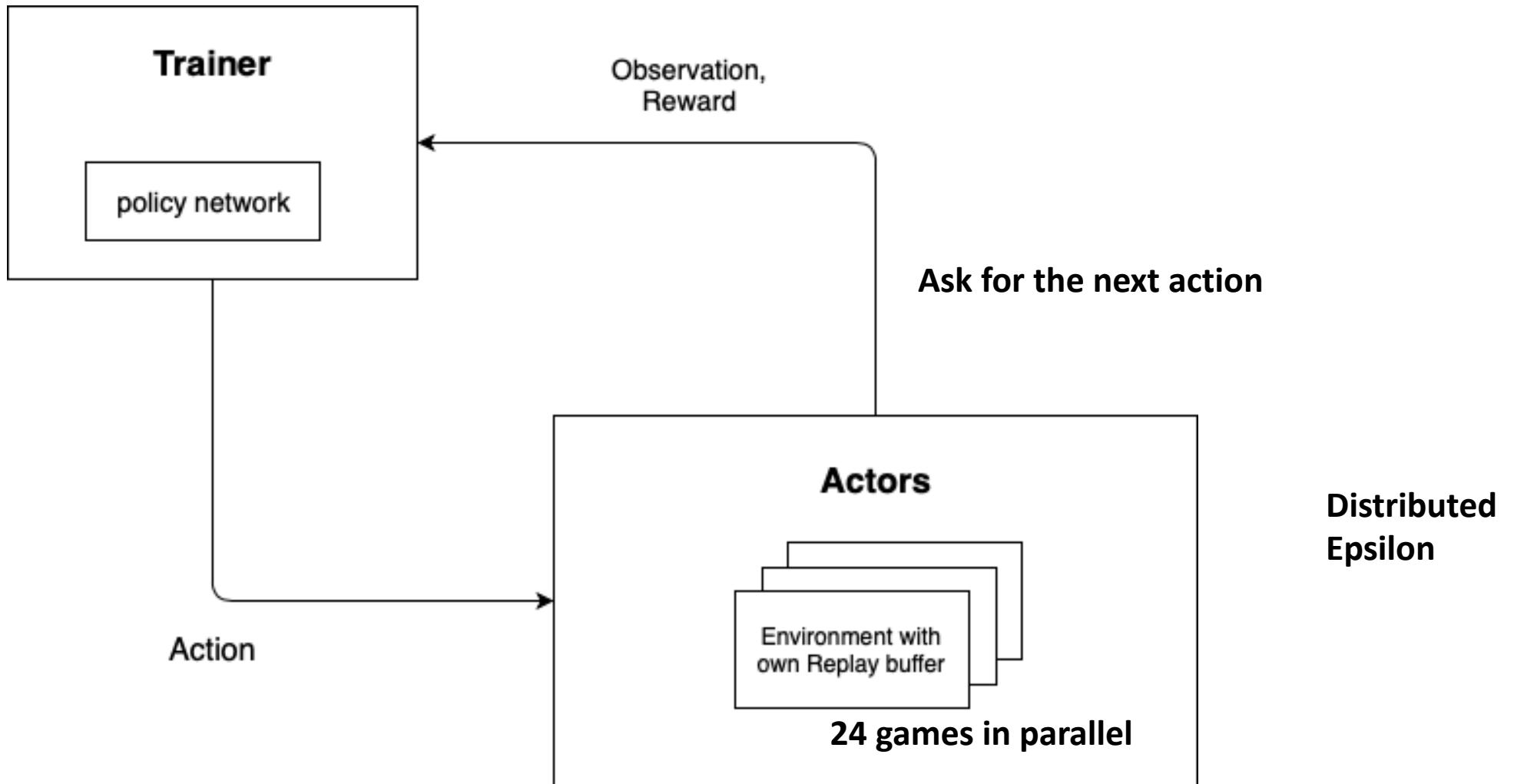


DQN algorithm



Usage of RLlib

Trains one central policy network



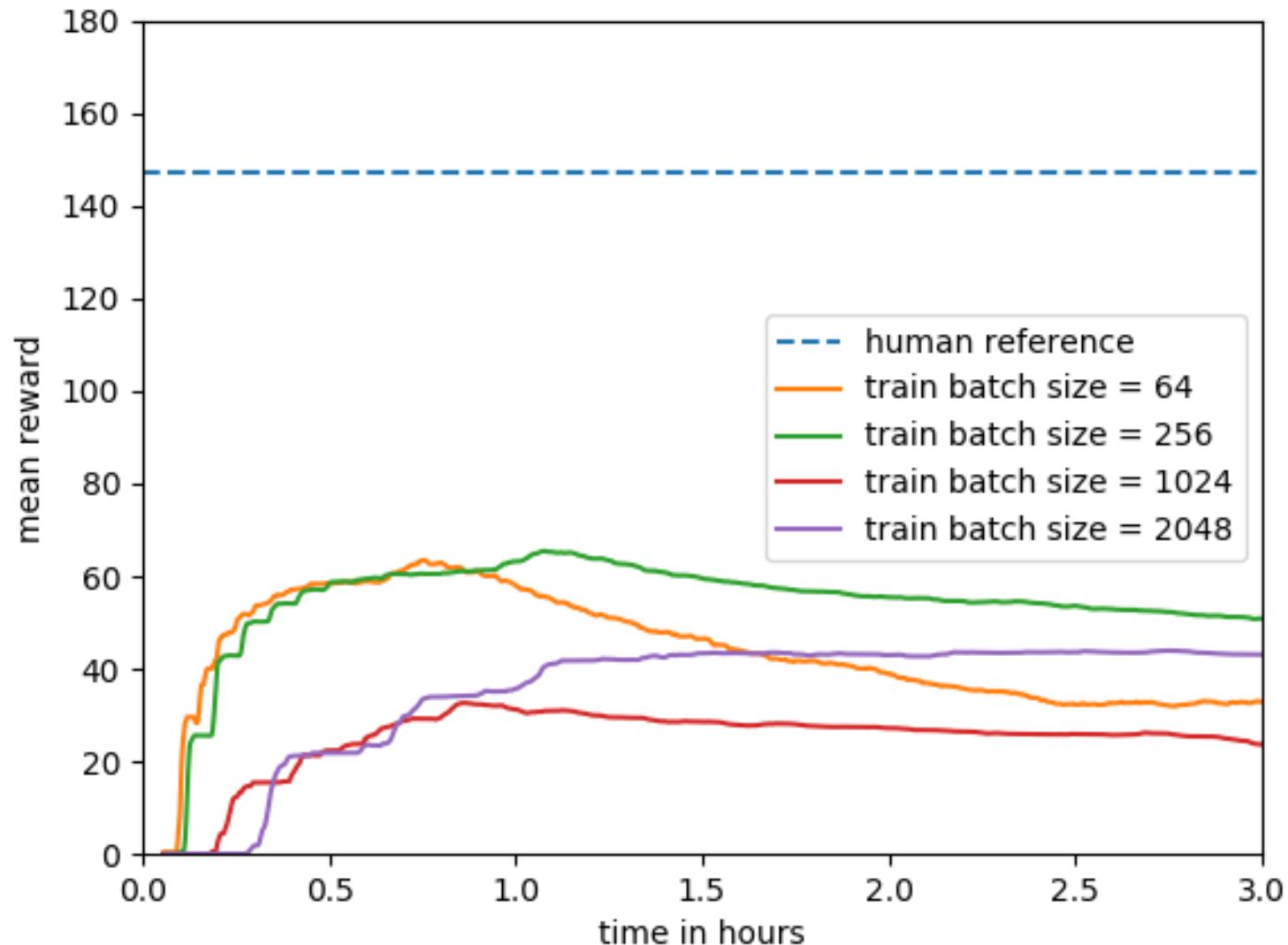
Results of DQN hyperparameter search

- Hyperparameter search for
 - Training batch size {64, 256, 1024, 2048}
 - Discount factor {0.99, 0.9, 0.8}
 - Learning rate {0.0001, 0.001, 0.01}

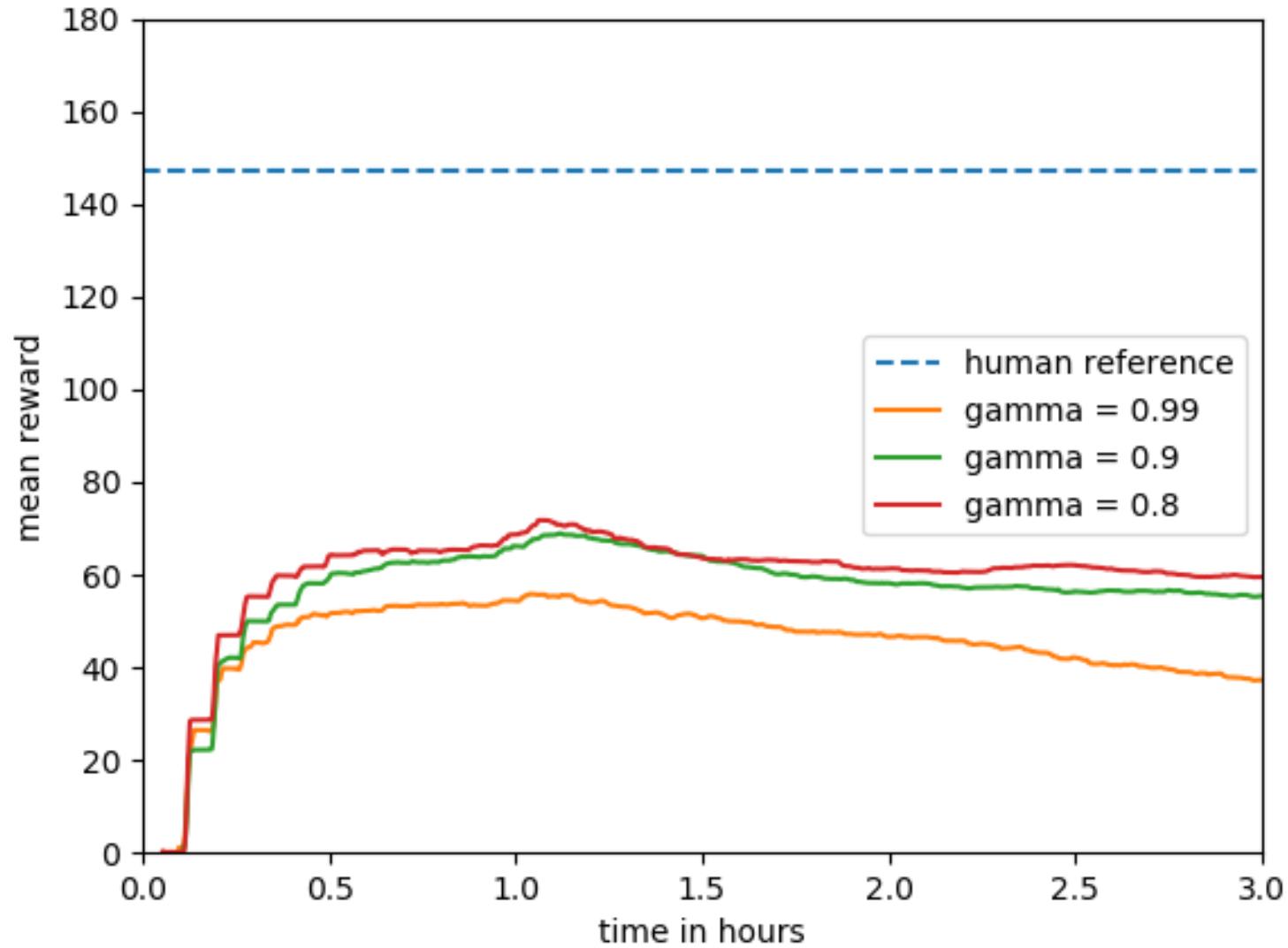
Results of DQN hyperparameter search

- Method:
 - 10 measurements per value per hyperparameter
 - 3 hours (24 parallel workers * 3 hours = 72 hours in serial) of training
 - Average values for comparison
 - Equally distributed ϵ between 0 and 0.4 for exploration

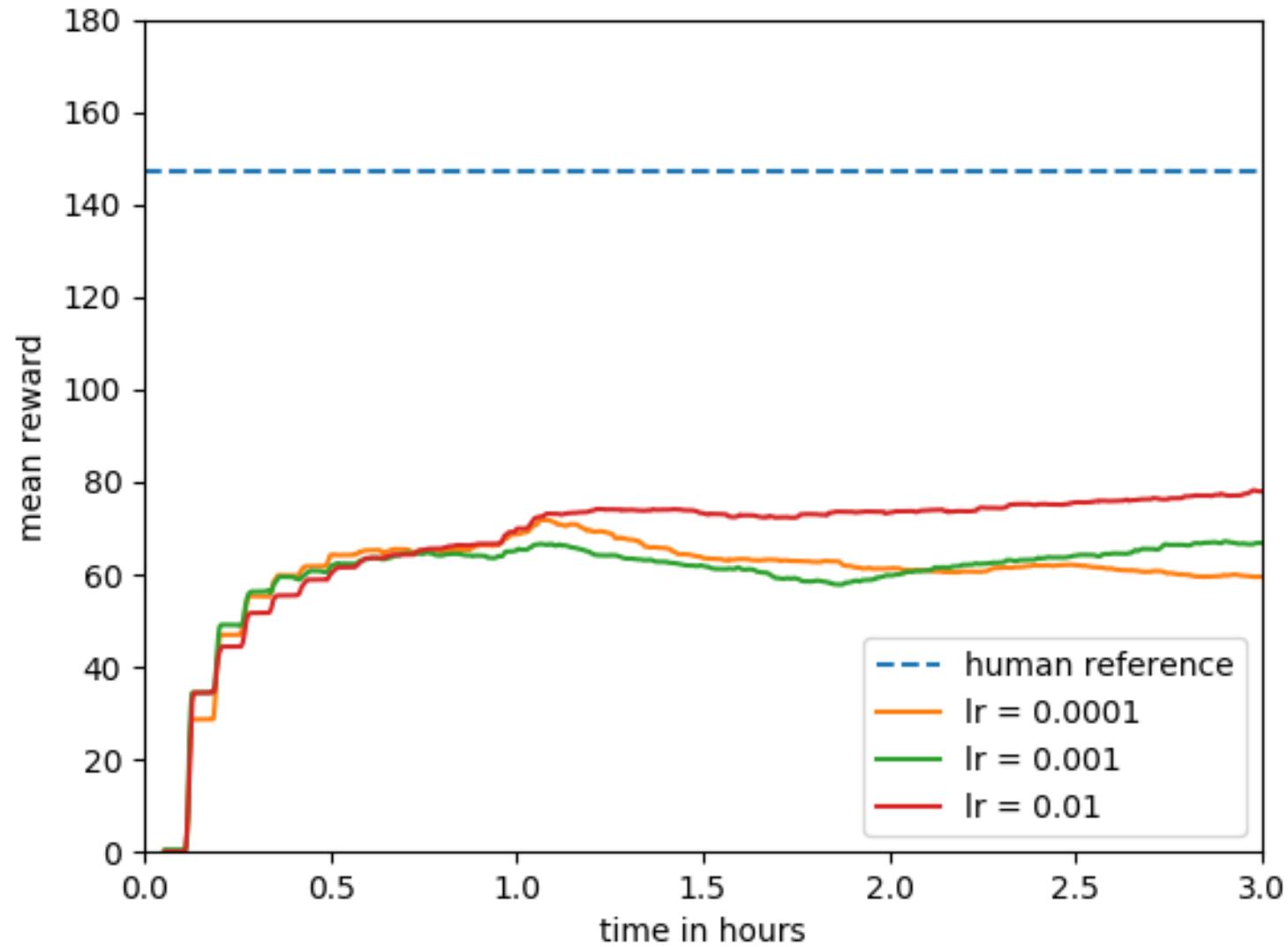
Results: Training batch size



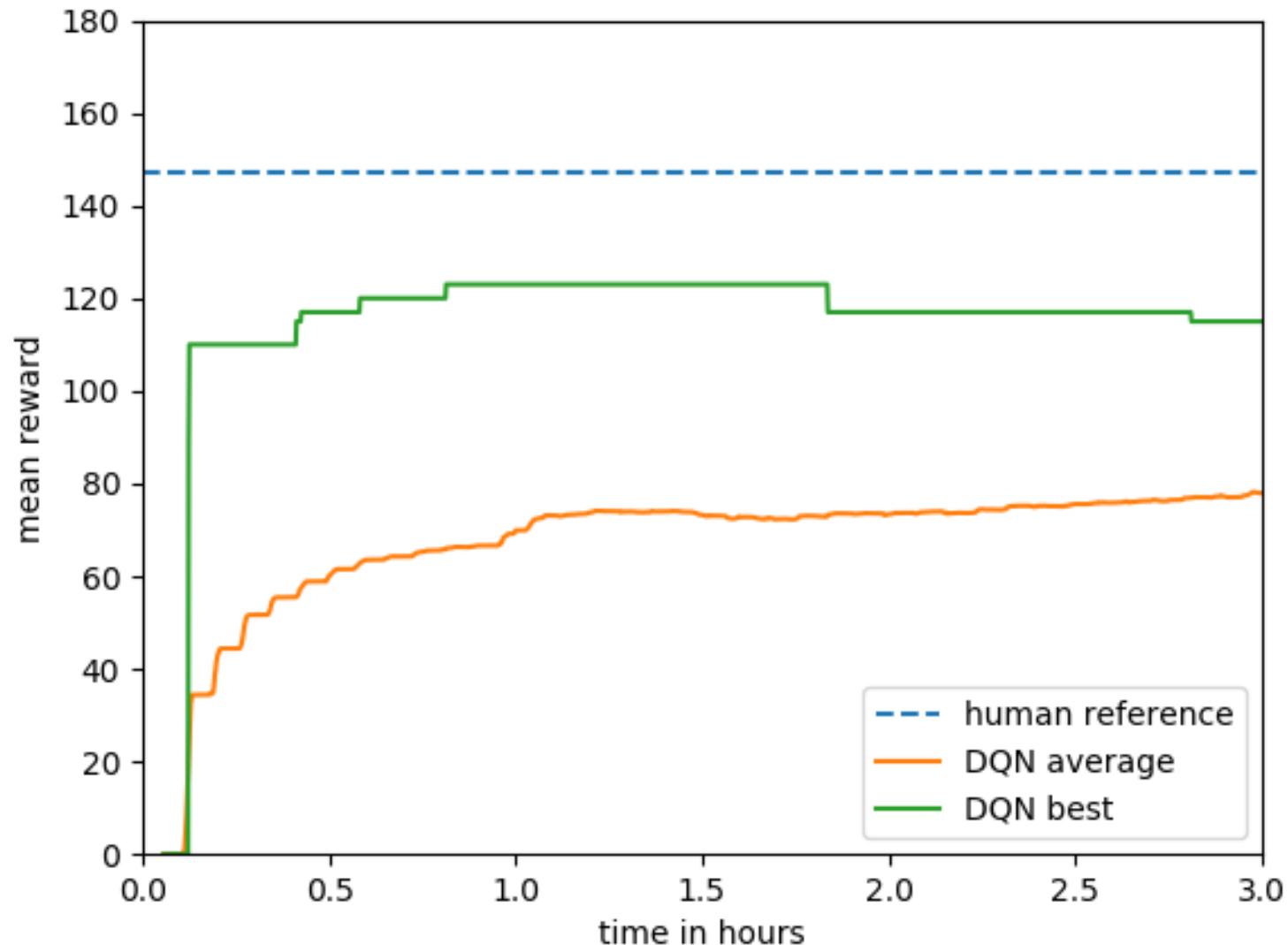
Results: Discount factor



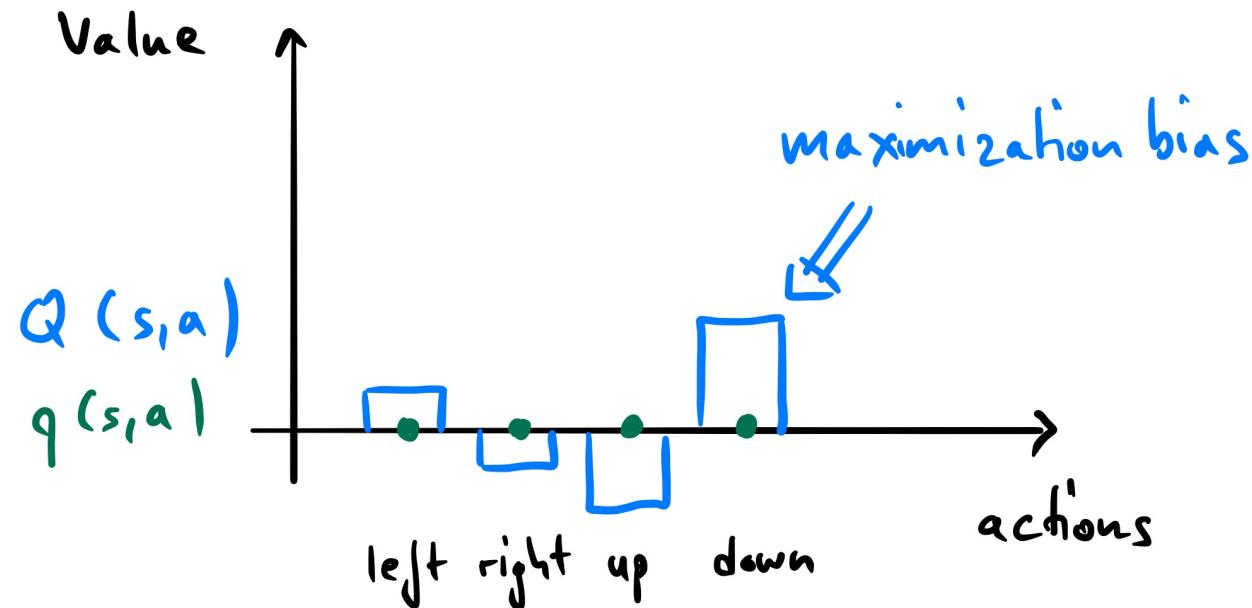
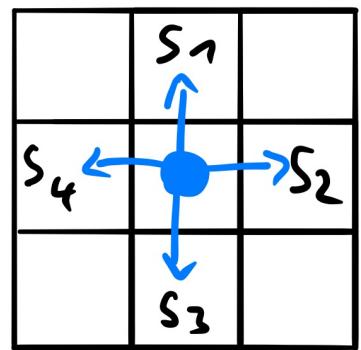
Results: Learning rate



Results: DQN



Double Learning



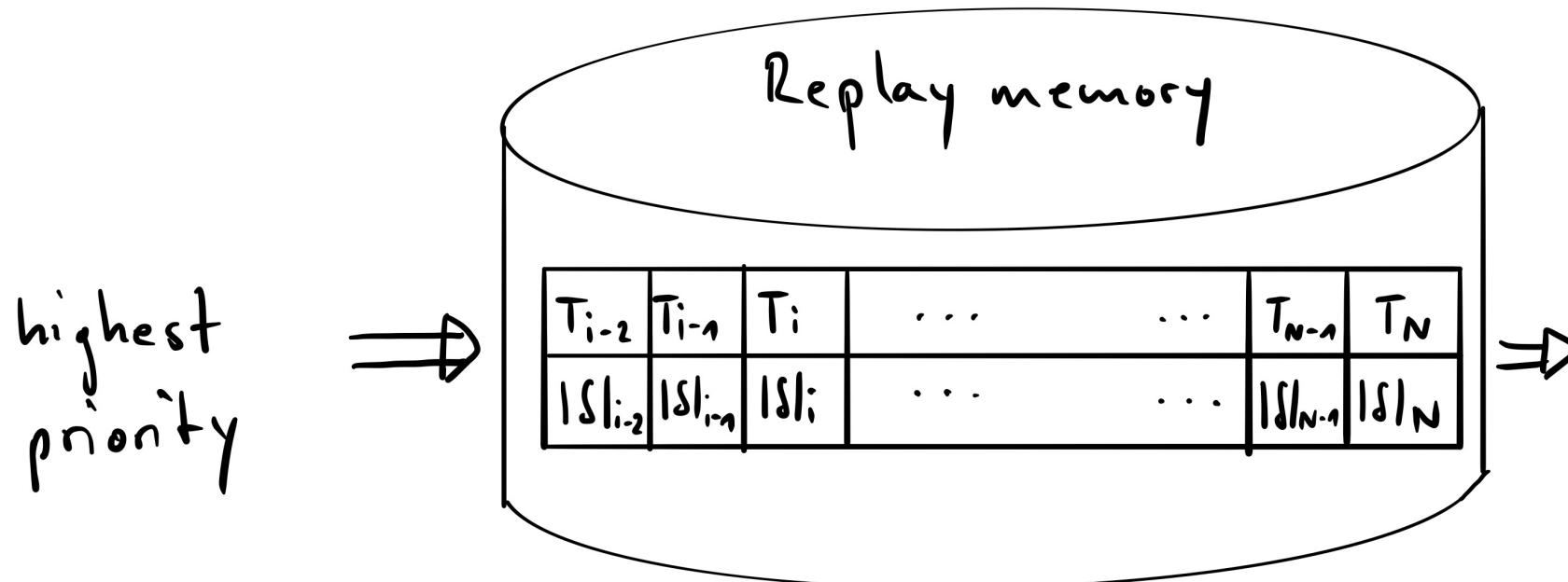
$$a_{\pi}^* = \arg \max_a Q(S_0, a)$$

$$V_{\pi} = Q(S_0, a)$$

$$a_{\pi}^* = \arg \max_a Q_1(S_0, a)$$

$$V_{\pi} = Q_2(S_0, a)$$

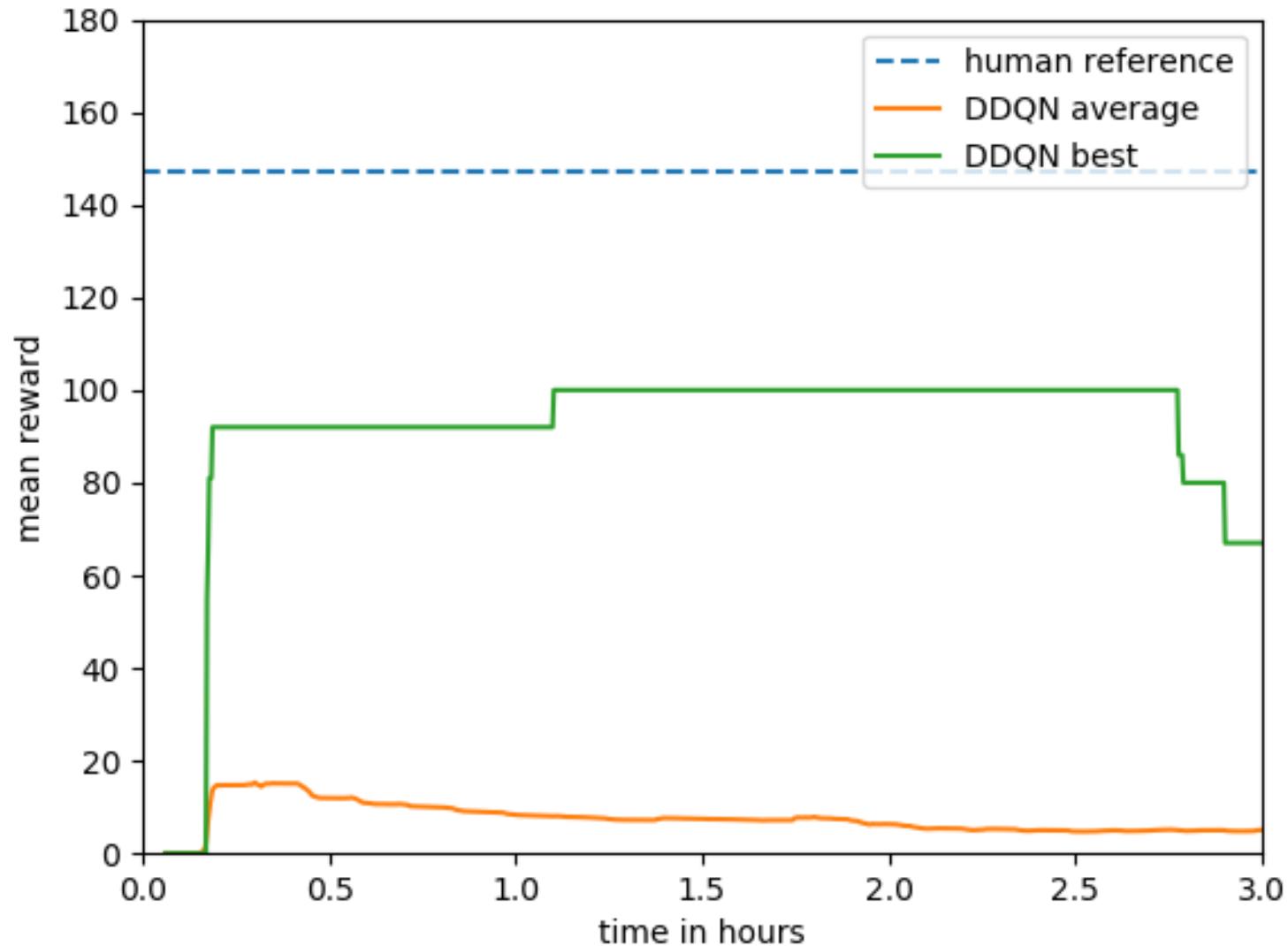
Prioritizing with TD-error



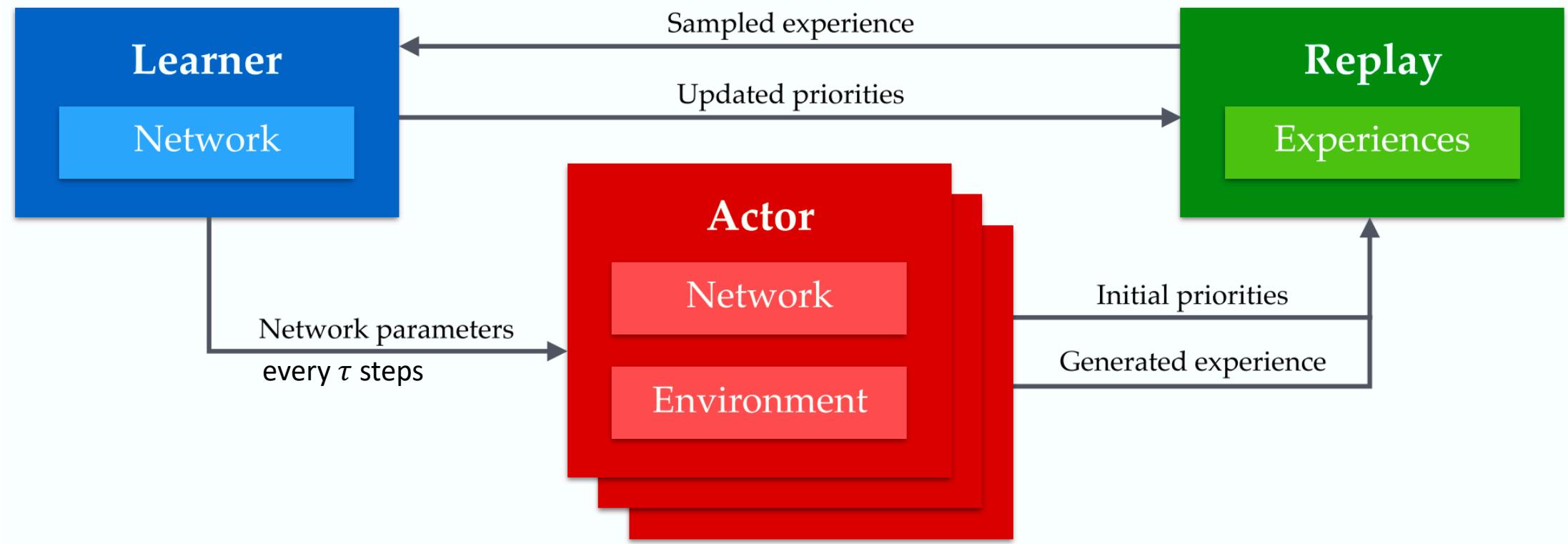
$$T = (S_t, A_t, R_{t+1}, S_{t+1})$$

$$|\delta| = |L_i(\theta_i)|$$

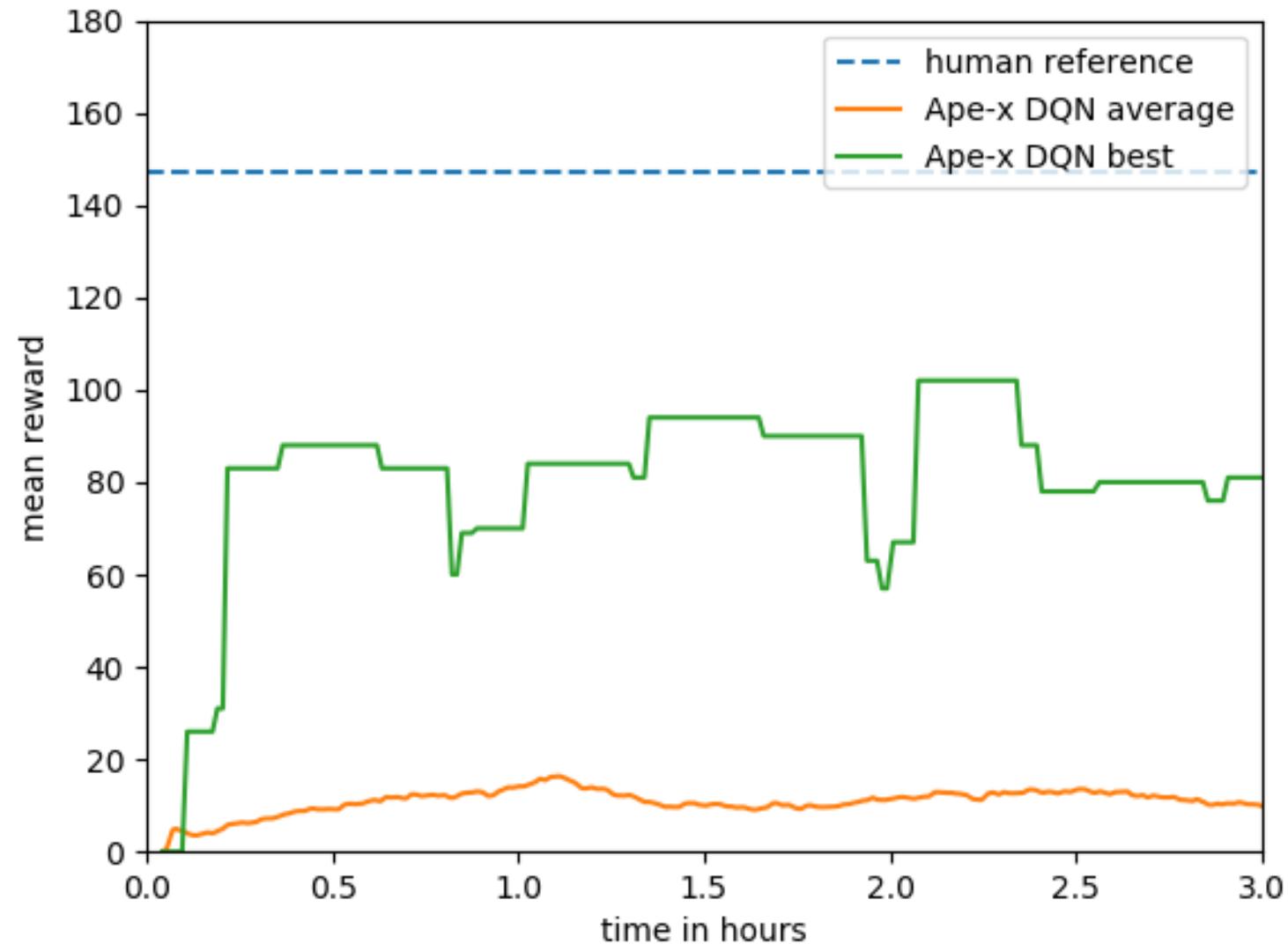
Results: DDQN



Ape-X Theory

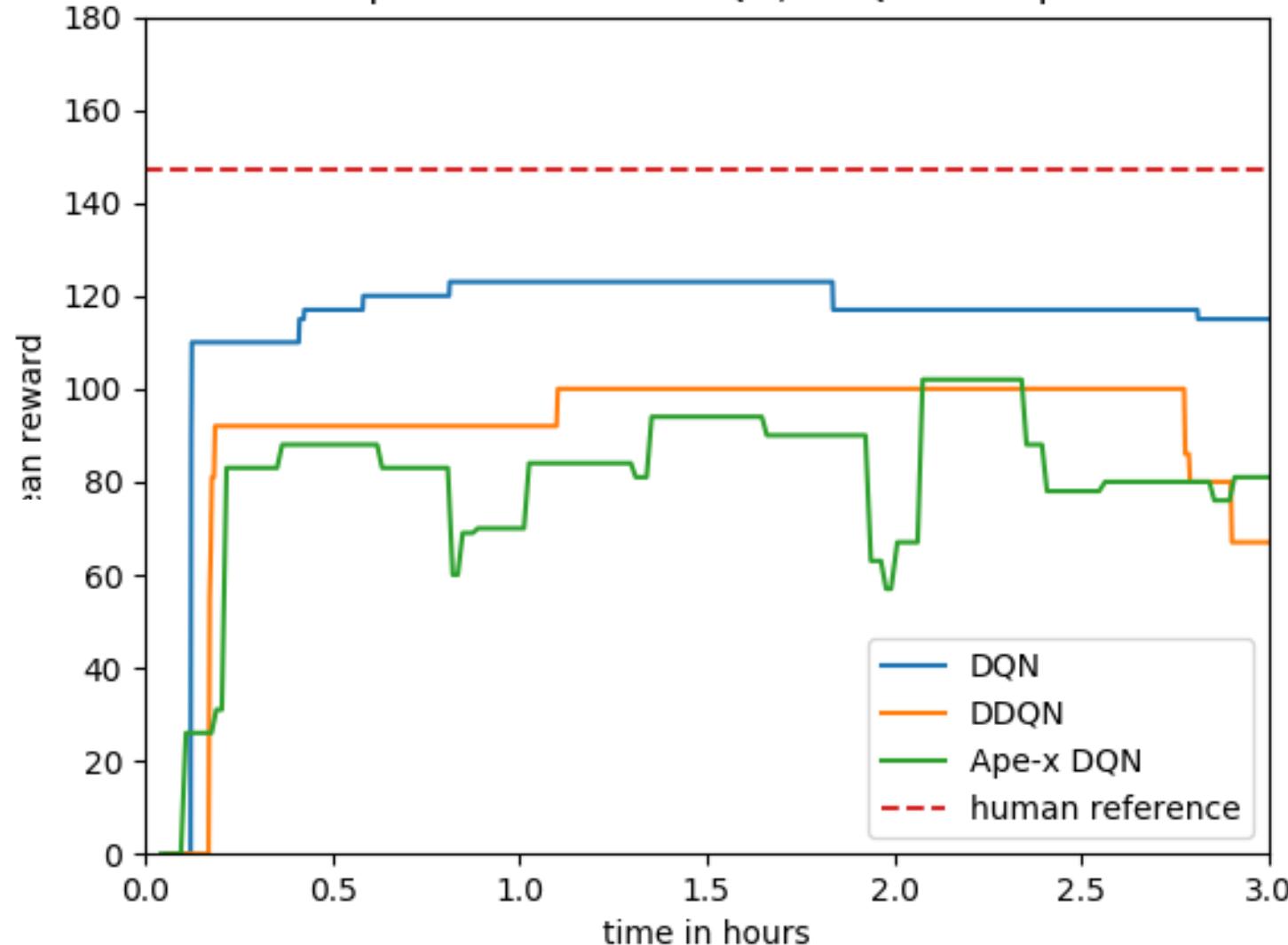


Results: Ape-X DQN



Best graphs

comparison of tuned DQN, DDQN and Ape-X



Future Work

- Spatial features
- Less atomic functions (less simplifications)
- DDQN and Ape-X DQN hyperparameter search for better mean results
- More than one mini-game

Lessons Learned

- First read papers and try to fully understand
- Do not waste time in doing side topics, focus on your topic
- Then implement it at your own to make sure you fully understood
- Only then - after that, use frameworks like RLlib
- Calculate a lot of extra time for troubleshooting

Questions

Thank you for your attention