

title
subtitle

XXX

University of Freiburg

February 6th, 2026

Our method: DQN

Extension 1: Double DQN

Extension 2: Multi-step Learning

Idea: Instead of using only the immediate reward, we use the reward of the next n steps.

- Instead of using the standard TD target:

$$R_{t+1} + \gamma \max_{a'} Q_\theta(s_{t+1}, a')$$

- We use the n-step return:

$$R_t^{(n)} + \gamma^{(n)} \max_{a'} Q_\theta(s_{t+n}, a')$$

- Where:

$$R_t^{(n)} = \sum_{k=0}^{n-1} \gamma^k R_{t+k+1} = R_{t+1} + \gamma R_{t+2} + \dots + \gamma^{n-1} R_{t+n}$$

Extension 3: Dueling DQN

Results

Conclusion