REINFORCE L-T vs. DQN L-T
$\delta d = 0.1$ $\epsilon d = 0.4$

- REINFORCE L-T
- DQN L-T

performance

context inference error