

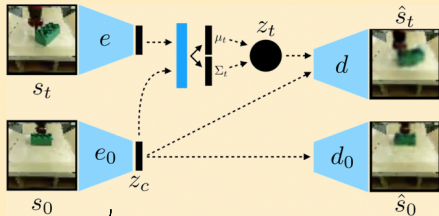


1. Collect random interaction samples

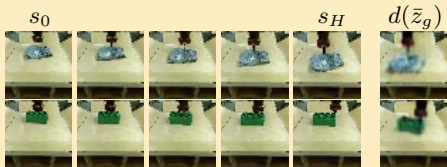


$\{\tau_1, \tau_2, \dots, \tau_N\}$

2. Train context-conditioned VAE

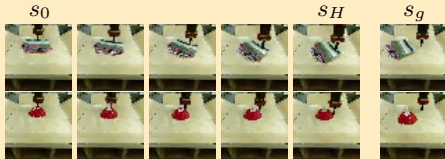


3. RL training: learn policy $\pi(\bar{z}_t, \bar{z}_g)$ to minimize latent distance to generated goal z_g



Real-world Pusher Training Rollouts

4. Test time: agent executes policy to reach human-provided goal image s_g



Test Rollouts (Unseen Objects)