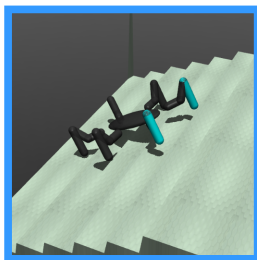
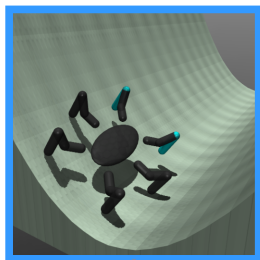


N environments



...



Reinforcement  
learning

$\pi_{\theta_1}^e$

$\pi_{\theta_2}^e$

...

$\pi_{\theta_n}^e$

Expert policies

