



● : Environment Sample ●/● : Model Sample P : Environment Transition M : Model Transition $D_{\text{env}}/D_{\text{model}}$: Replay Buffer