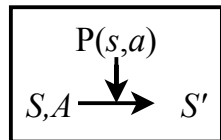


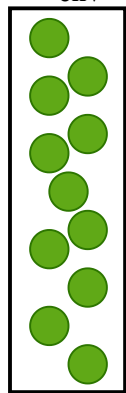
## Environment



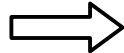
Sample



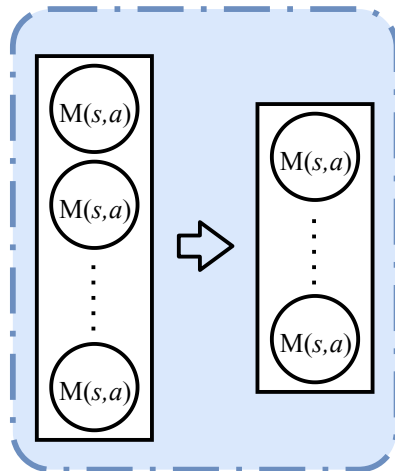
$D_{\text{env}}$



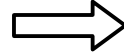
Train



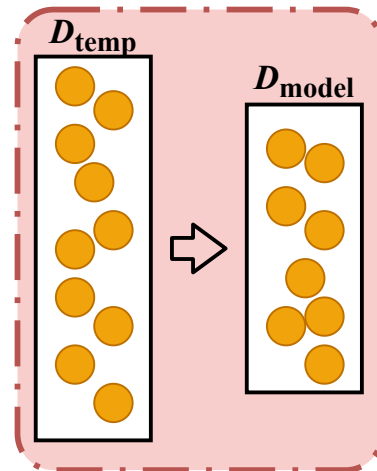
## Model Dropout



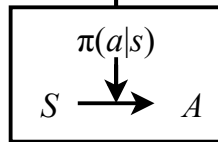
Rollout



## Rollout Dropout



Policy



Policy Optimization

 : Environment Sample     : Model Sample     $P(s,a)$  : Environment Transition     $M(s,a)$  : Model Transition     $D_{\text{env}}/D_{\text{model}}$  : Replay Buffer