

# Controller

Blue: Controller  
Green: Environment  
Orange: Observations

$a_{t-1}$

$\hat{s}_t$

$a_t$

$\hat{s}_{t+1}$

Post-Action  
Observation

$\hat{f}_t$   
 $o_t$   
 $f_t$   
 $t$

$\hat{f}_{t+1}$   
 $o_{t+1}$   
 $f_{t+1}$   
 $t+1$

Time

$t-1$

$s_{t-1}$

$P(\cdot | s_{t-1}, a_{t-1})$

$s_t$

$P(\cdot | s_t, a_t)$

$s_{t+1}$

$s_{t+1} \sim P(\cdot | s_t, a_t)$

# Environment