

state  
 $S_t$

reward  
 $R_t$

$R_{t+1}$

$S_{t+1}$

Environment

Agent

action  
 $A_t$

