

$b_0$ : initial belief  
 $s$ : state  
 $a$ : action  
 $r$ : reward  
 $o$ : observation

