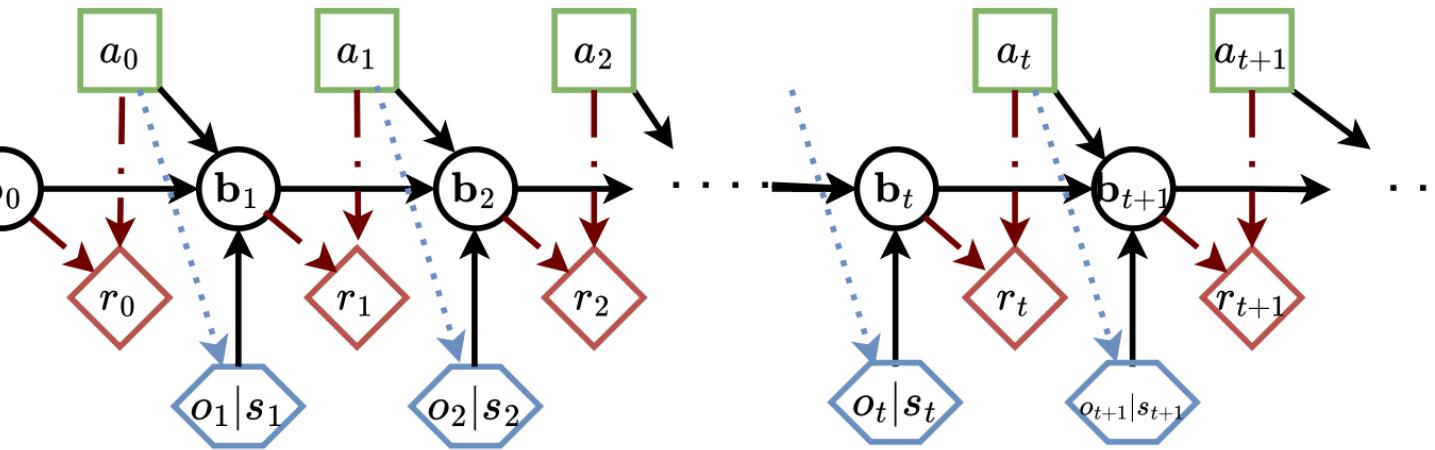


$b_t$	<b>Belief:</b>	$b_t \in \mathbb{B} \subseteq \mathbb{R}^{ \mathbb{B} }$
$a_t$	<b>Action:</b>	$a_t \in \mathbb{A}$
$r_t$	<b>Reward:</b>	$r_t \in \mathbb{R}$
$o_t   s_t$	<b>Observation:</b>	$o_t \in \mathbb{O}$



→ **Transition Model:**

$$\begin{aligned} & P(\mathbf{b}_{t+1} | \mathbf{b}_t, a_t) \\ &= \int_{o_{t+1} \in \mathbb{O}} P(o_{t+1} | \mathbf{b}_t, a_t) do_{t+1} \\ &= \sum_{o_{t+1} \in \mathbb{O}} P(o_{t+1} | \mathbf{b}_t, a_t) \end{aligned}$$

→ → **Reward Model:**

$$\begin{aligned} & r(a_t, \mathbf{b}_t) \\ &= \int_{s_t \in \mathbb{S}} r(a_t, s_t) b_t(s_t) ds_t \\ &= \sum_{s_t \in \mathbb{S}} r(a_t, s_t) b_t(s_t) \end{aligned}$$