

\bullet b_t

Belief: $b_t \in \mathbb{B} \subseteq \mathbb{R}^{|\mathbb{B}|}$

\square a_t

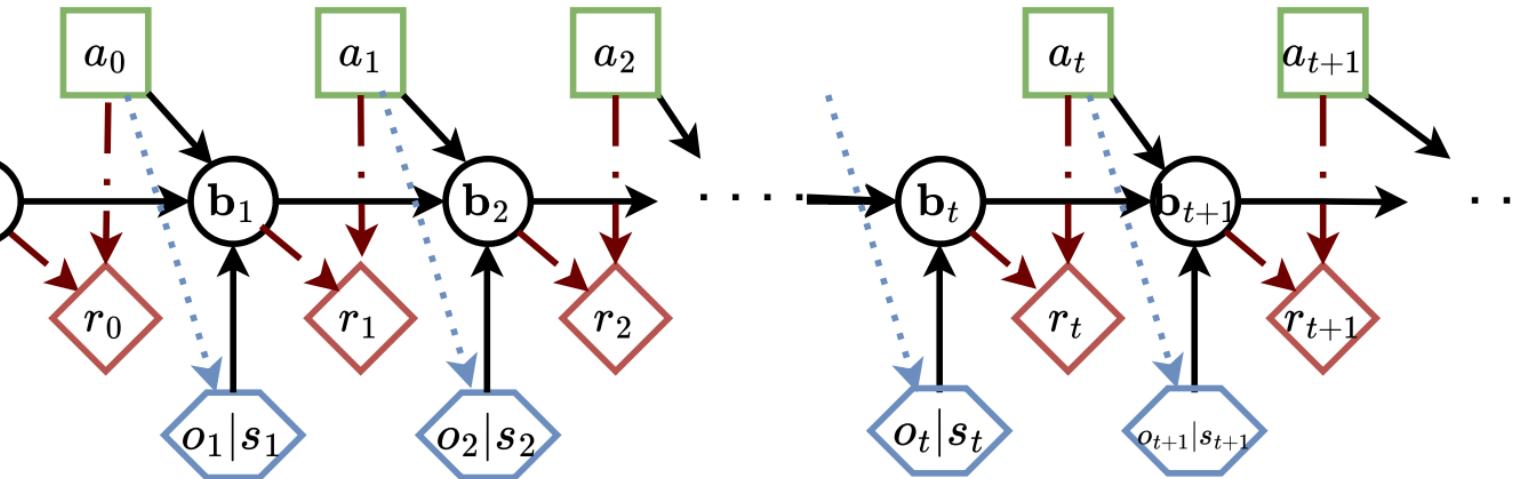
Action: $a_t \in \mathbb{A}$

\diamond r_t

Reward: $r_t \in \mathbb{R}$

\lozenge $o_t | s_t$

Observation: $o_t \in \mathbb{O}$



Transition Model: $P(b_{t+1}|b_t, a_t) = \int_{o_{t+1} \in \mathbb{O}} P(o_{t+1}|b_t, a_t) do_{t+1}$

$$P(b_{t+1}|b_t, a_t) = \sum_{o_{t+1} \in \mathbb{O}} P(o_{t+1}|b_t, a_t)$$



Reward Model: $r(a_t, b_t) = \int_{s_t \in \mathbb{S}} r(a_t, s_t) b_t(s_t)$

$$r(a_t, b_t) = \sum_{s_t \in \mathbb{S}} r(a_t, s_t) b_t(s_t)$$