

$s_t$

$a_t$

$r_t$

**State:**

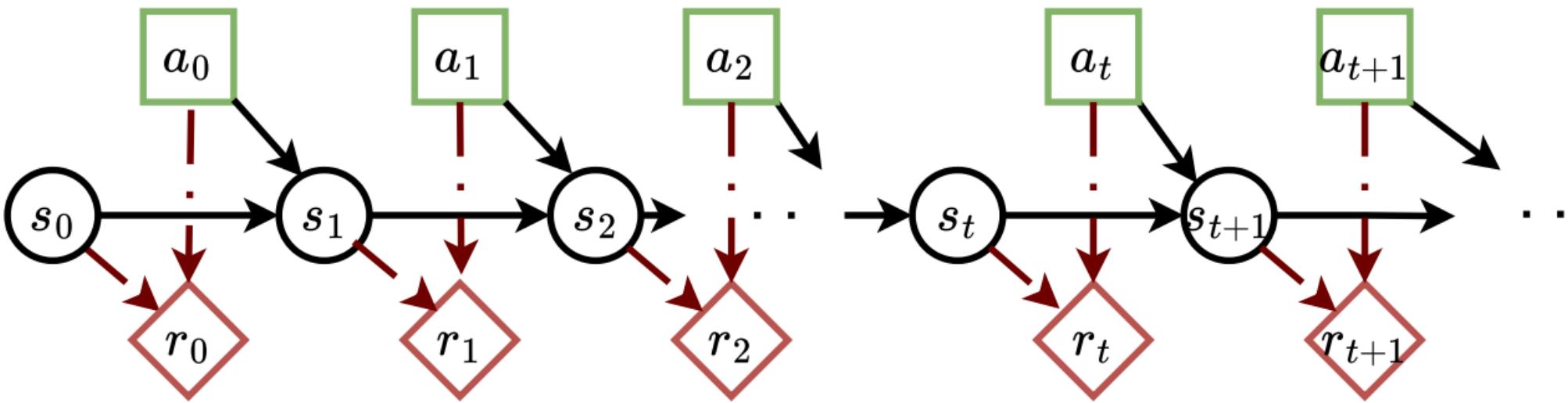
$s_t \in \mathbb{S}$

**Action:**

$a_t \in \mathbb{A}$

**Reward:**

$r_t \in \mathbb{R}$



→ **Transition Model:**  $\mathbf{T} = P(s_{t+1}|s_t, a_t)$

→ → **Reward Model:**  $\mathbf{R} = P(r_t|s_t, a_t)$

or  $\mathbf{R} = R(s_t, a_t)$  for deterministic reward model