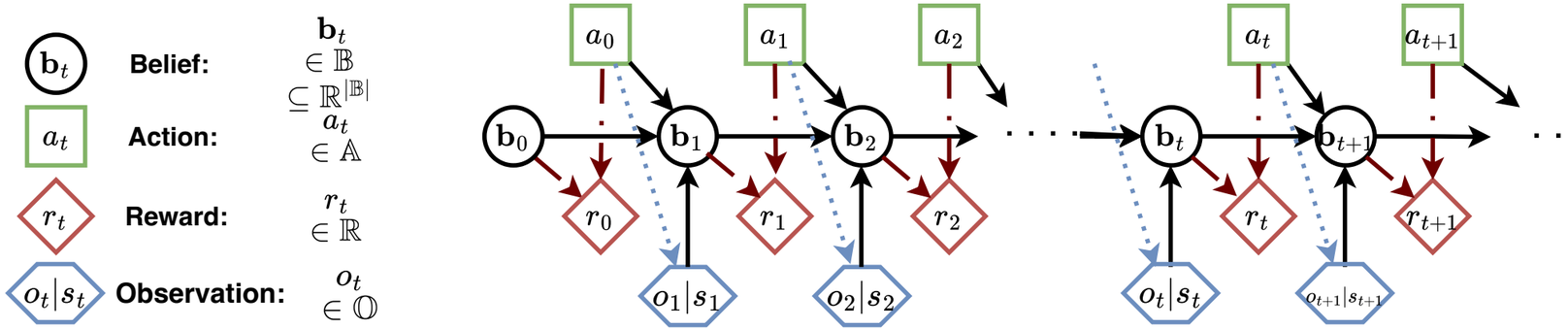

Transition Model:

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

$$P(\mathbf{b}_{t+1}|\mathbf{b}_t, a_t) = \sum_{o_{t+1}\in\mathbb{O}} P(o_{t+1}|\mathbf{b}_t, a_t)$$

Reward Model:

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

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