



Environment



Action (a_t)

Reward (r_t, a_t)

Current state
(s_t)

Next state
(s_{t+1})

Agent
(SNOW node/BS)