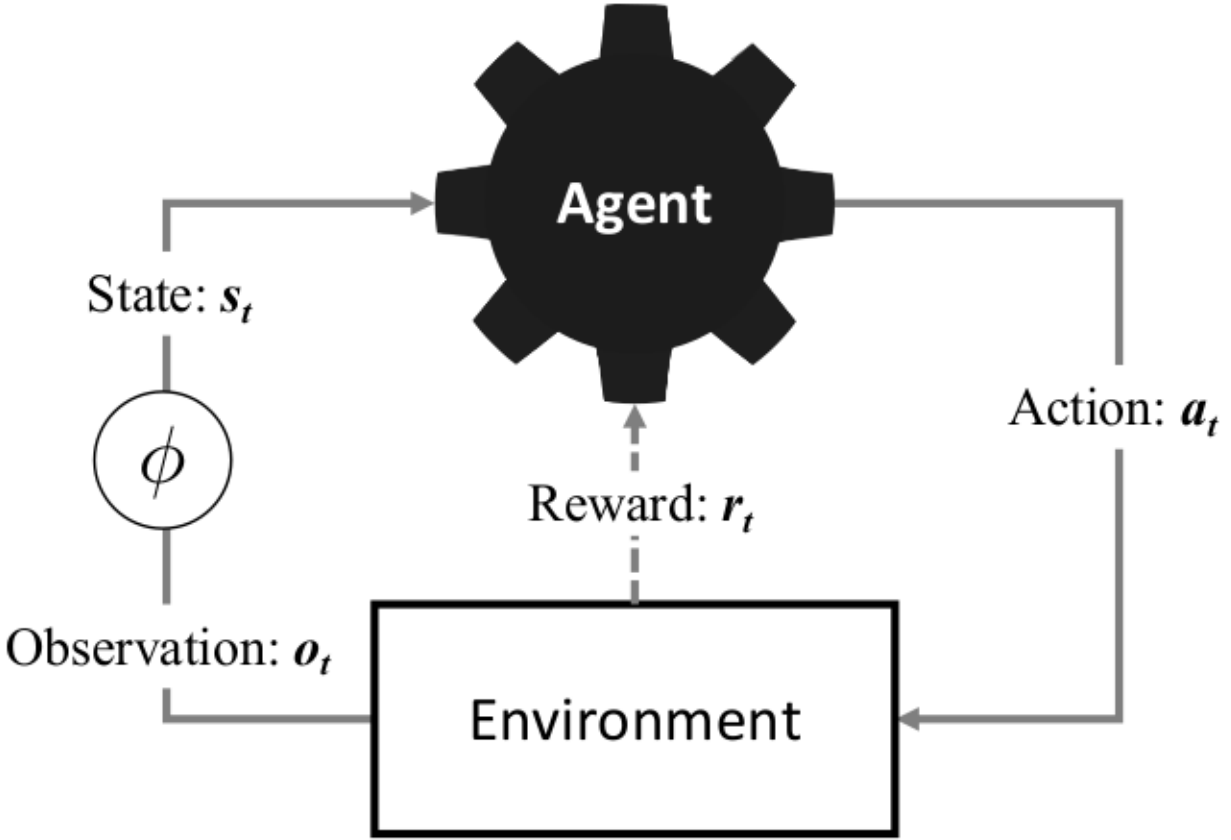
Agent
State: s_t
ϕ
Observation: o_t
Reward: r_t
Action: a_t
Environment