Environment

action

*At*

*Rt+1*

*St+1*

Agent

reward

*Rt*

state

*St*