Agents

Environments

action A_t

reward R_t R_{t+1}

state S_t S_{t+1}