



Perception

estimation

state
 S_t

reward
 R_t

R_{t+1}

S_{t+1}

Agent

Planning

Environment

control reference

action
 A_t

Control

ground truth

control value

