

# 外部环境

数据  $s$

数据  $s$

学习算法

知识表征

内部损失函数:  $J$

内部动力学模型:  $M$

决策模型:  
 $d(s | J, M)$

动作  $a$

机器人行为系统

