belief state update
with observations,
leverage context

action selection
maximize expected value

execute action
"black box"
receive observations

Time