

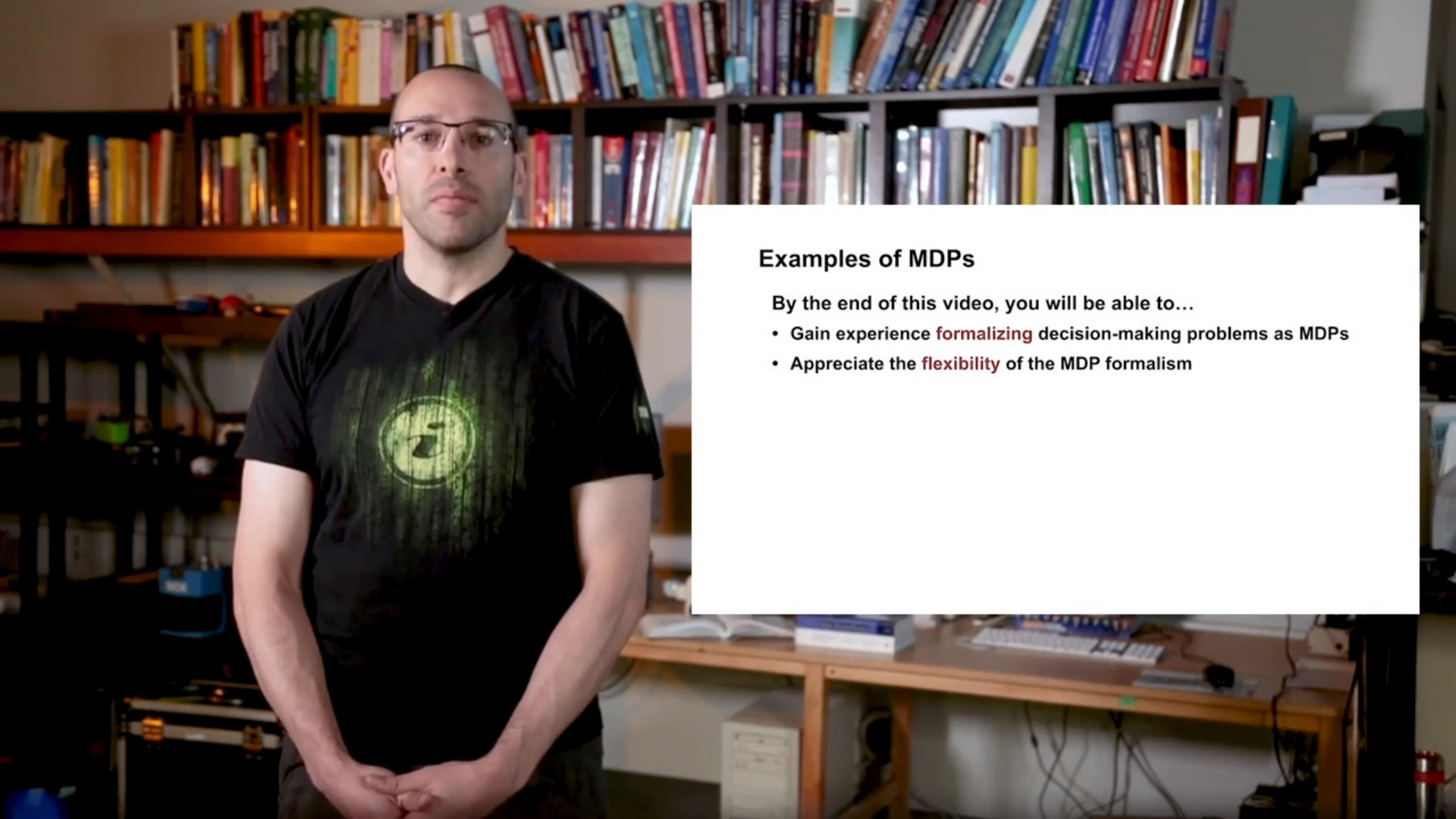


## Examples of MDPs

By the end of this video, you will be able to...

- Gain experience **formalizing** decision-making problems as MDPs



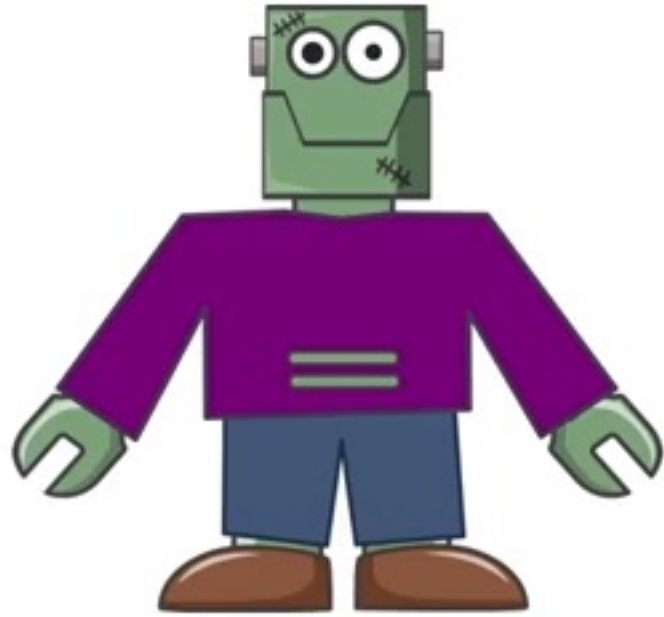


## Examples of MDPs

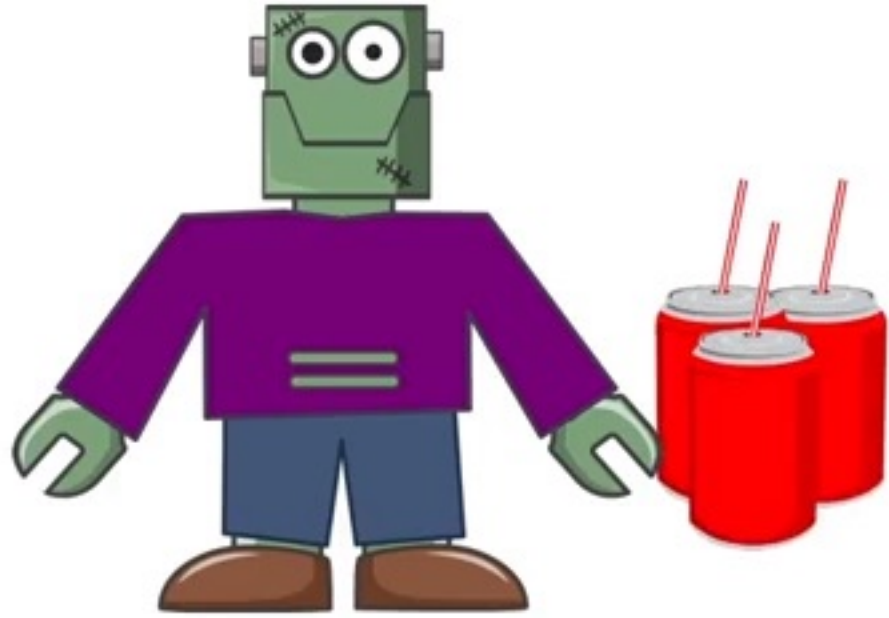
By the end of this video, you will be able to...

- Gain experience **formalizing** decision-making problems as MDPs
- Appreciate the **flexibility** of the MDP formalism

# Recycling Robot



# Recycling Robot



# Recycling Robot



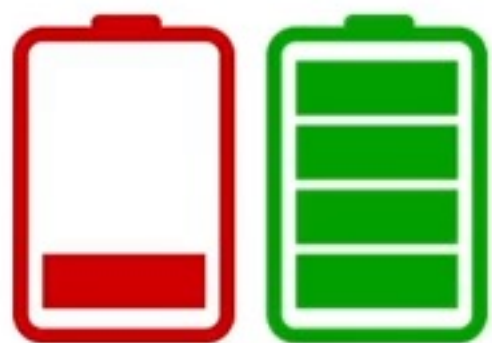


# Recycling Robot



$\mathcal{S} = \{\text{low, high}\}$

# Recycling Robot

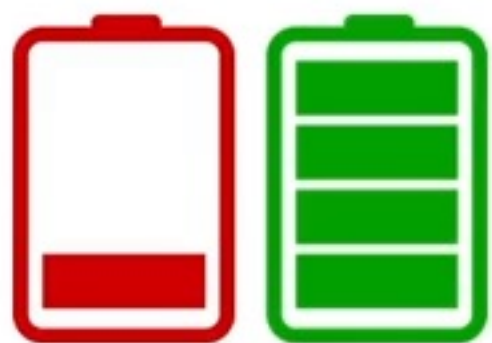


$\mathcal{S} = \{\text{low}, \text{high}\}$

$\mathcal{A}(\text{low})$

$\mathcal{A}(\text{high})$

# Recycling Robot



$\mathcal{S} = \{\text{low}, \text{high}\}$

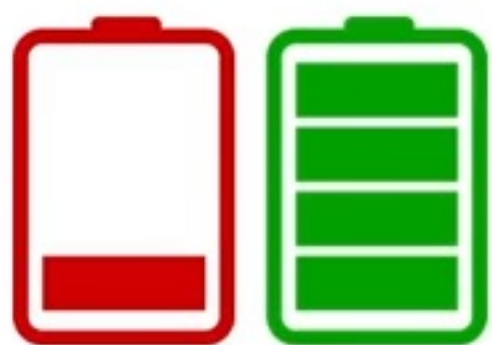


$\mathcal{A}(\text{low}) = \{\text{search}$

$\mathcal{A}(\text{high}) = \{\text{search}$



# Recycling Robot



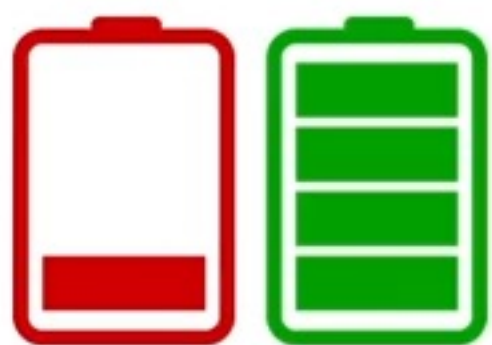
$\mathcal{S} = \{\text{low}, \text{high}\}$



$\mathcal{A}(\text{low}) = \{\text{search}, \text{wait}\}$

$\mathcal{A}(\text{high}) = \{\text{search}, \text{wait}\}$

# Recycling Robot



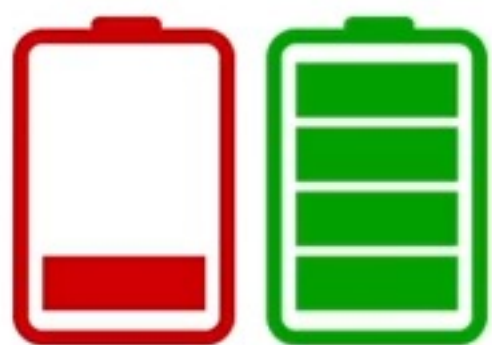
$\mathcal{S} = \{\text{low}, \text{high}\}$



$\mathcal{A}(\text{low}) = \{\text{search}, \text{wait}, \text{recharge}\}$

$\mathcal{A}(\text{high}) = \{\text{search}, \text{wait}, \text{recharge}\}$

# Recycling Robot



$\mathcal{S} = \{\text{low, high}\}$



$\mathcal{A}(\text{low}) = \{\text{search, wait, recharge}\}$

$\mathcal{A}(\text{high}) = \{\text{search, wait}\}$

# Dynamics of the Recycling Robot

**high**

**low**



# Dynamics of the Recycling Robot

search with energy level high may  
reduce the energy level to low



high

low

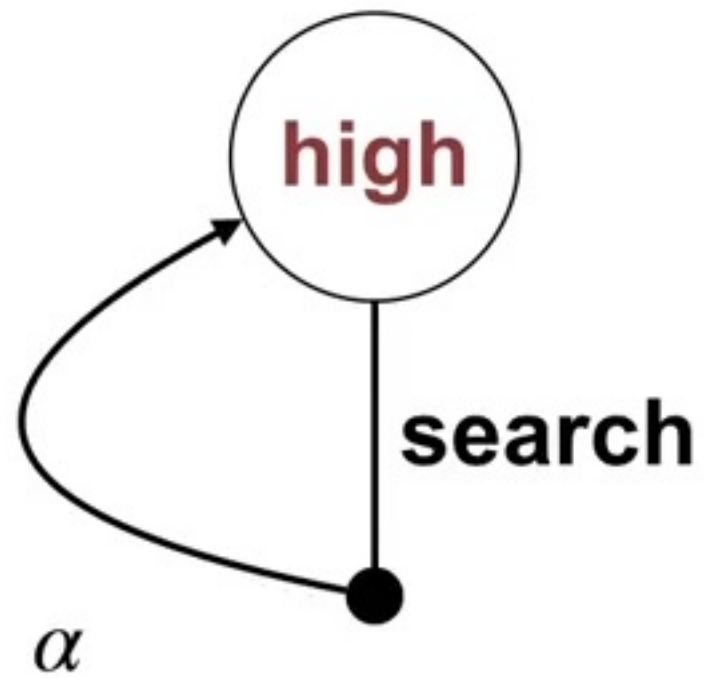
# Dynamics of the Recycling Robot

search with energy level high may  
reduce the energy level to low



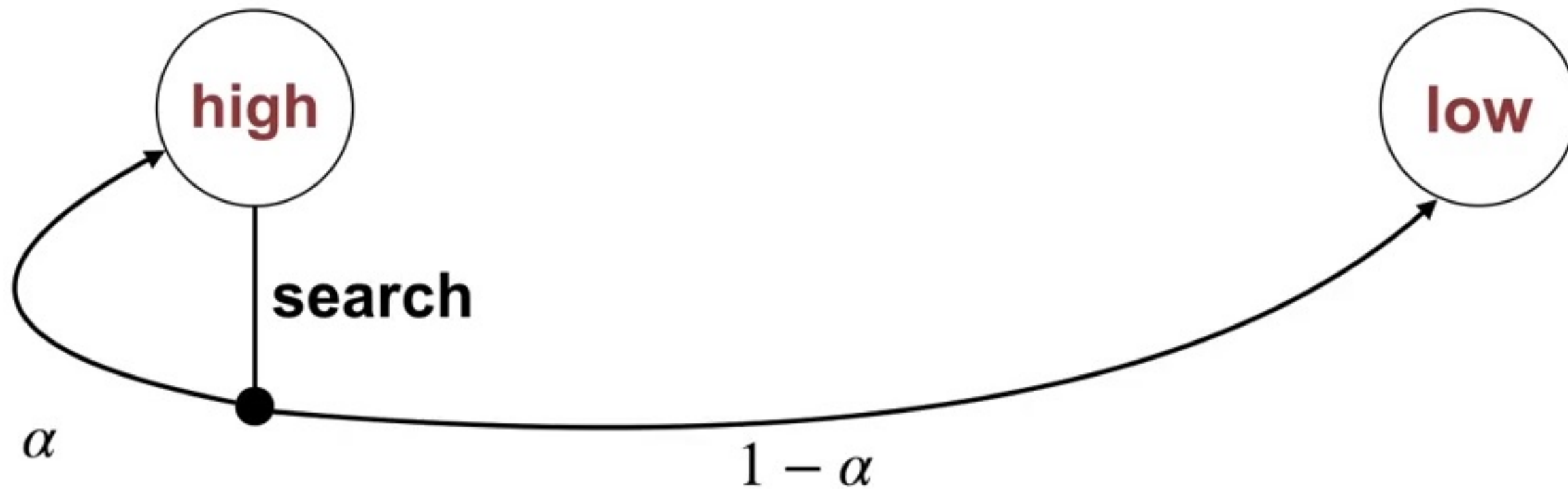
# Dynamics of the Recycling Robot

search with energy level high may  
reduce the energy level to low



# Dynamics of the Recycling Robot

search with energy level high may  
reduce the energy level to low





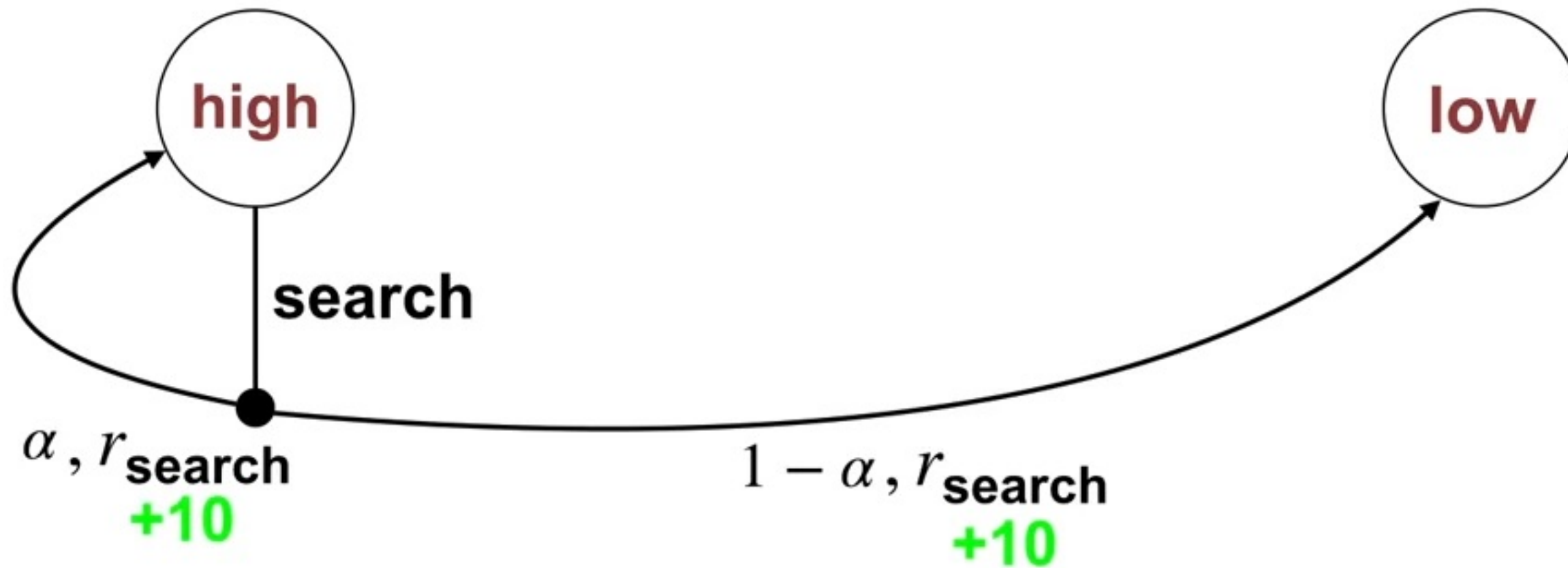
# Dynamics of the Recycling Robot

search with energy level high may  
reduce the energy level to low



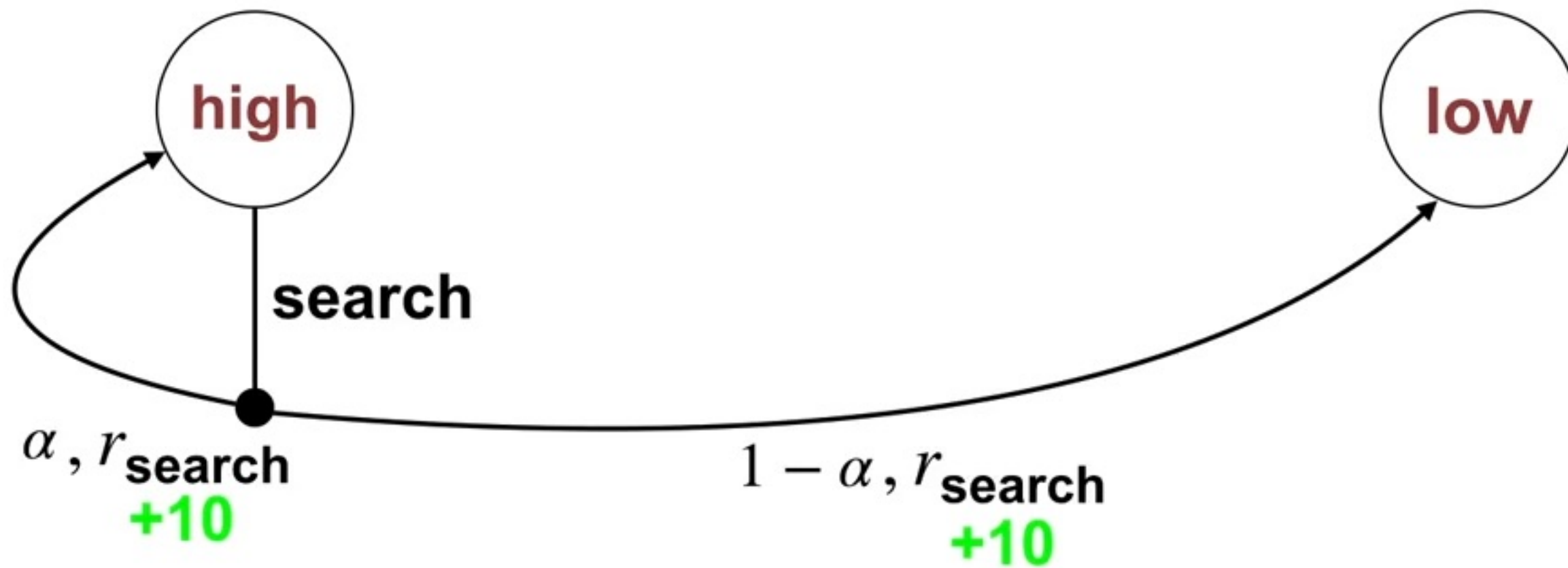
# Dynamics of the Recycling Robot

search with energy level high may  
reduce the energy level to low



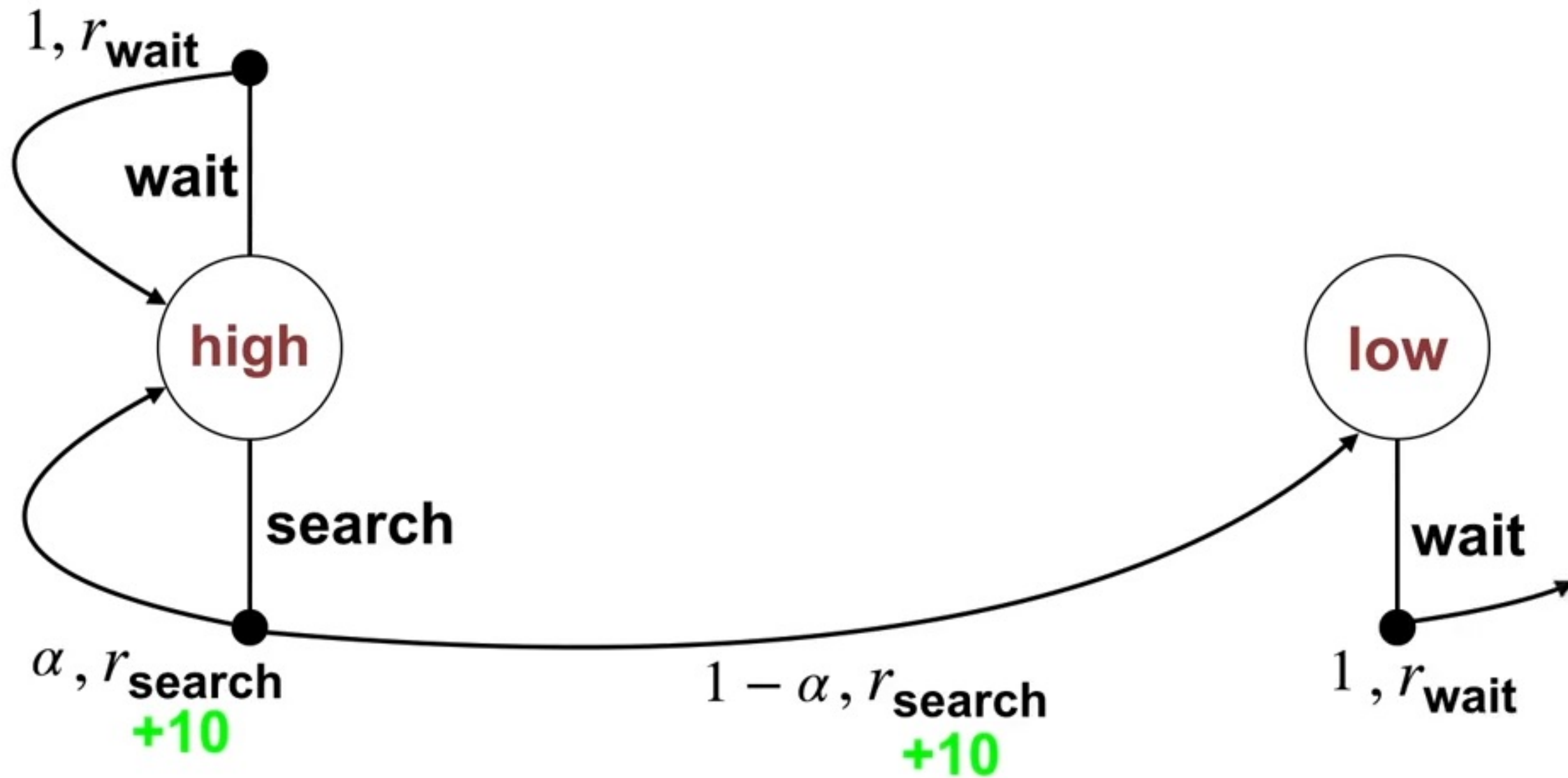
# Dynamics of the Recycling Robot

waiting for cans does not drain the battery



# Dynamics of the Recycling Robot

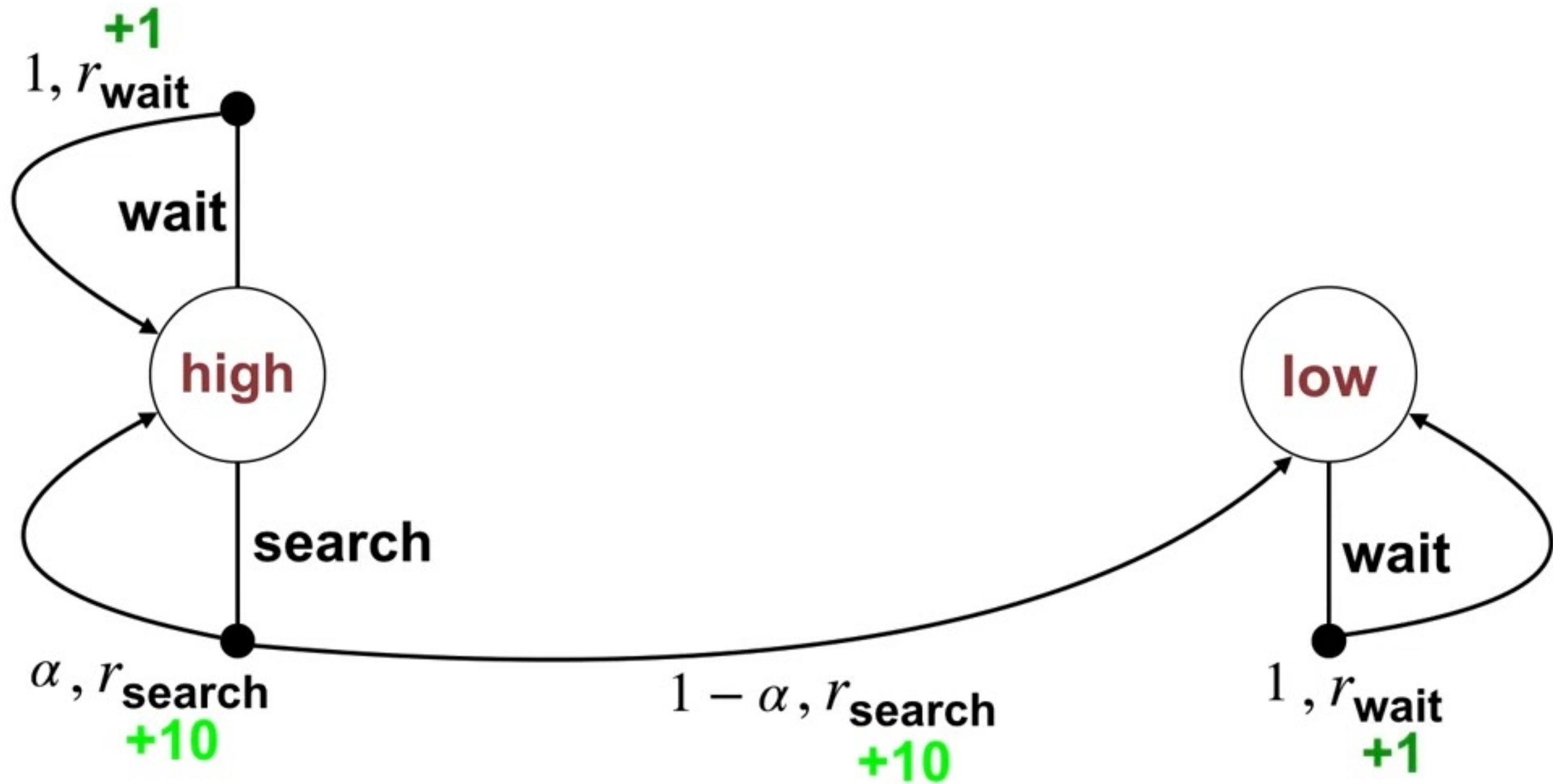
waiting for cans does not drain the battery





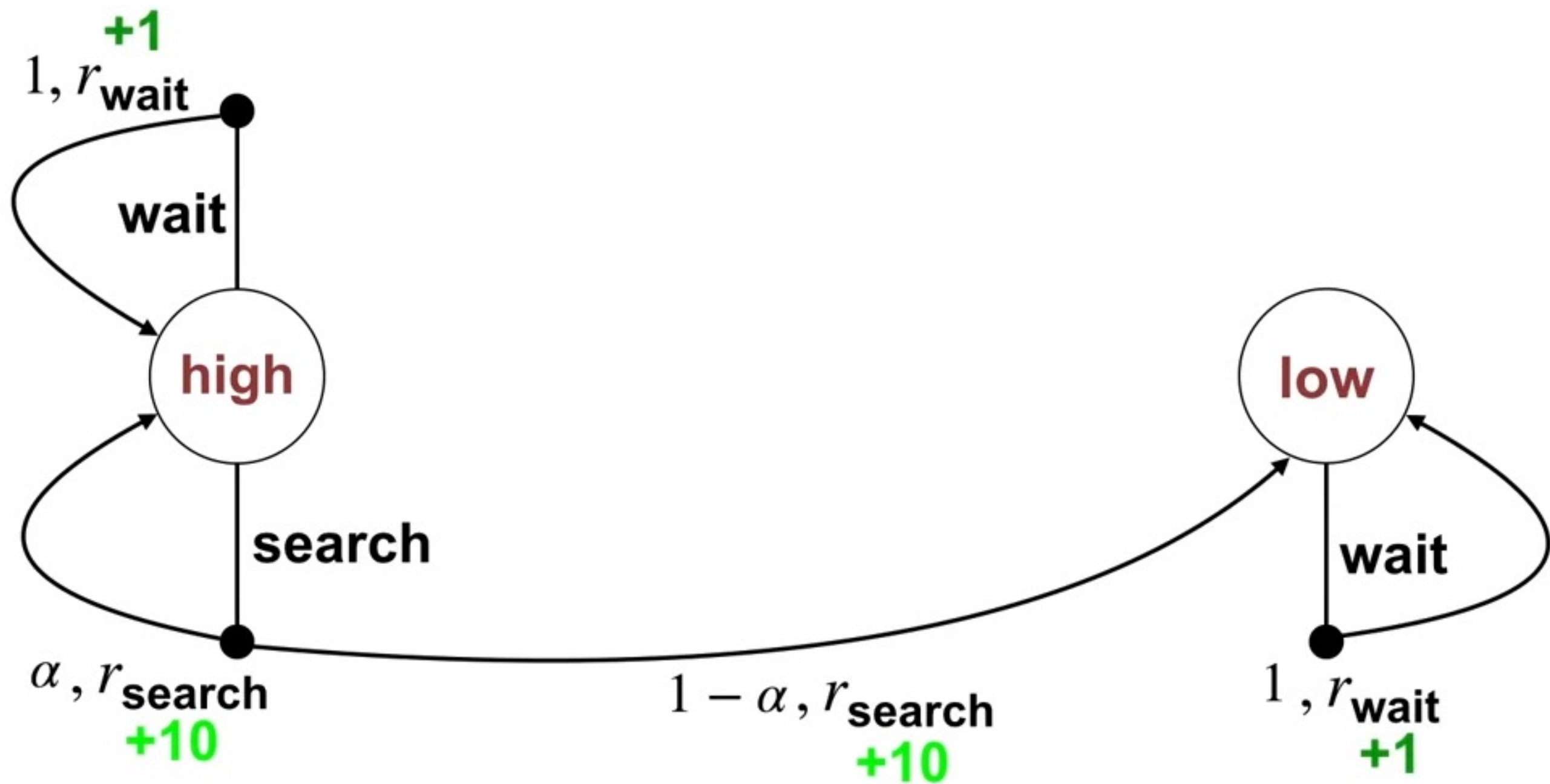
# Dynamics of the Recycling Robot

waiting for cans does not drain the battery



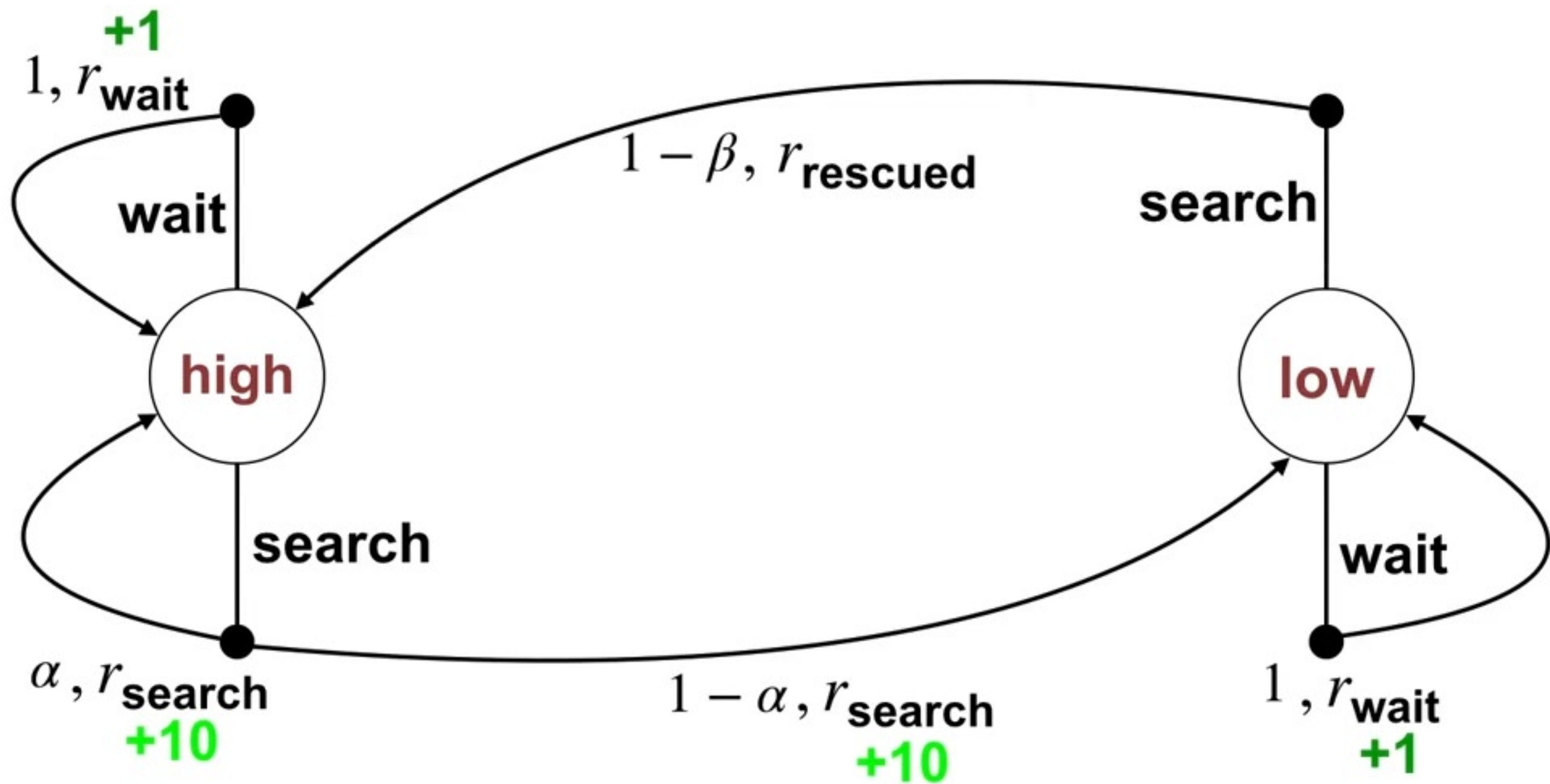
# Dynamics of the Recycling Robot

search with energy level low may deplete the battery  
and the robot must be rescued



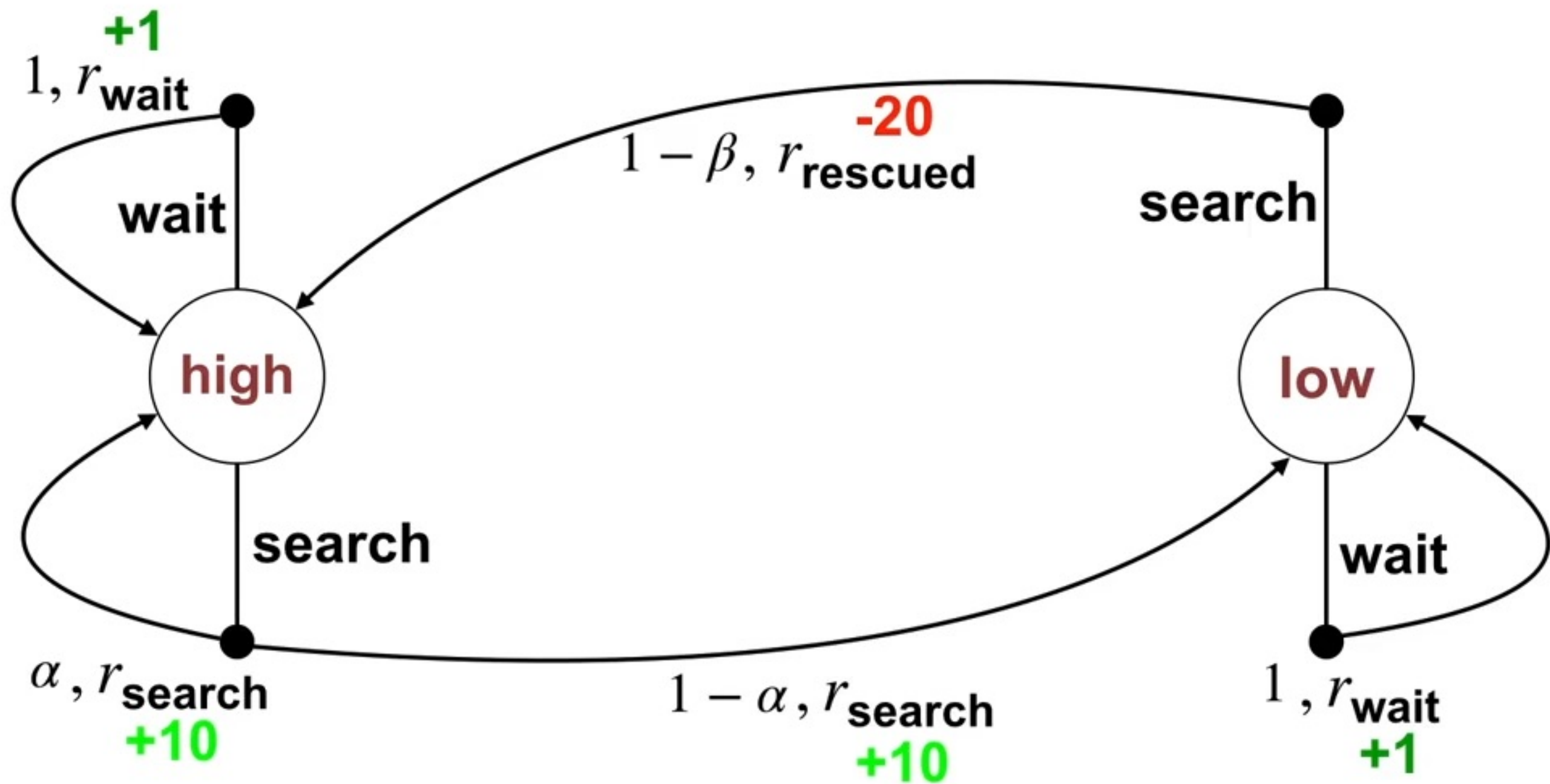
# Dynamics of the Recycling Robot

search with energy level low may deplete the battery and the robot must be rescued



# Dynamics of the Recycling Robot

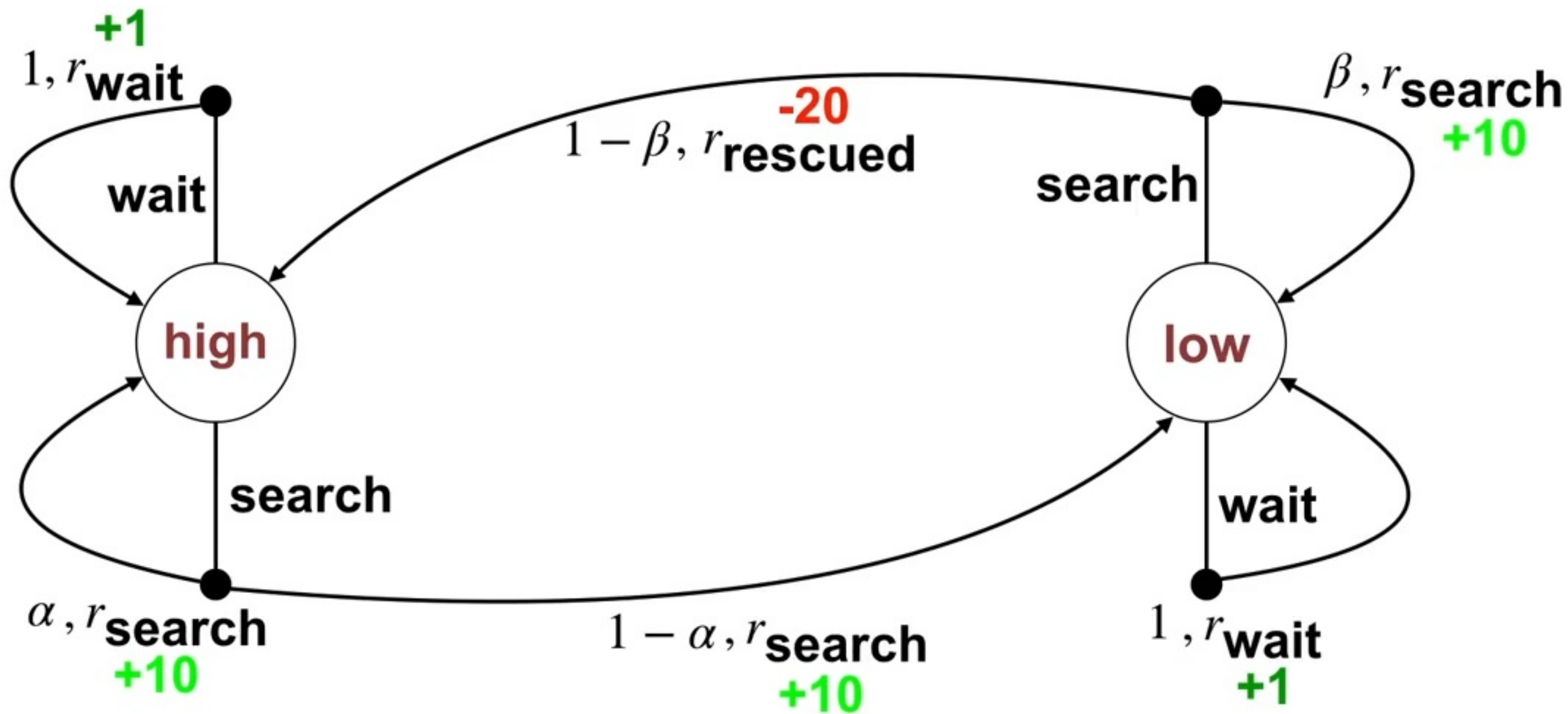
search with energy level low may deplete the battery  
and the robot must be rescued





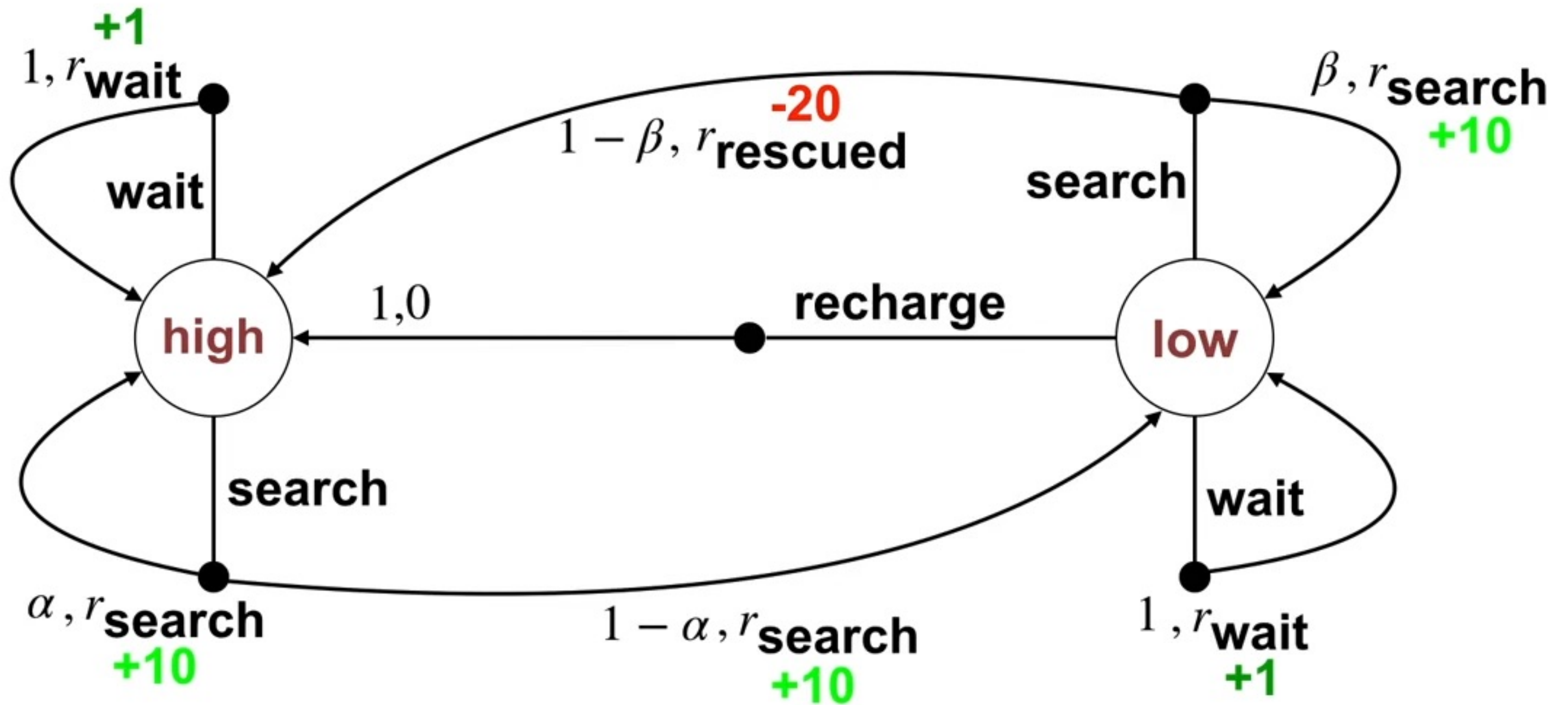
# Dynamics of the Recycling Robot

search with energy level low may deplete the battery  
and the robot must be rescued



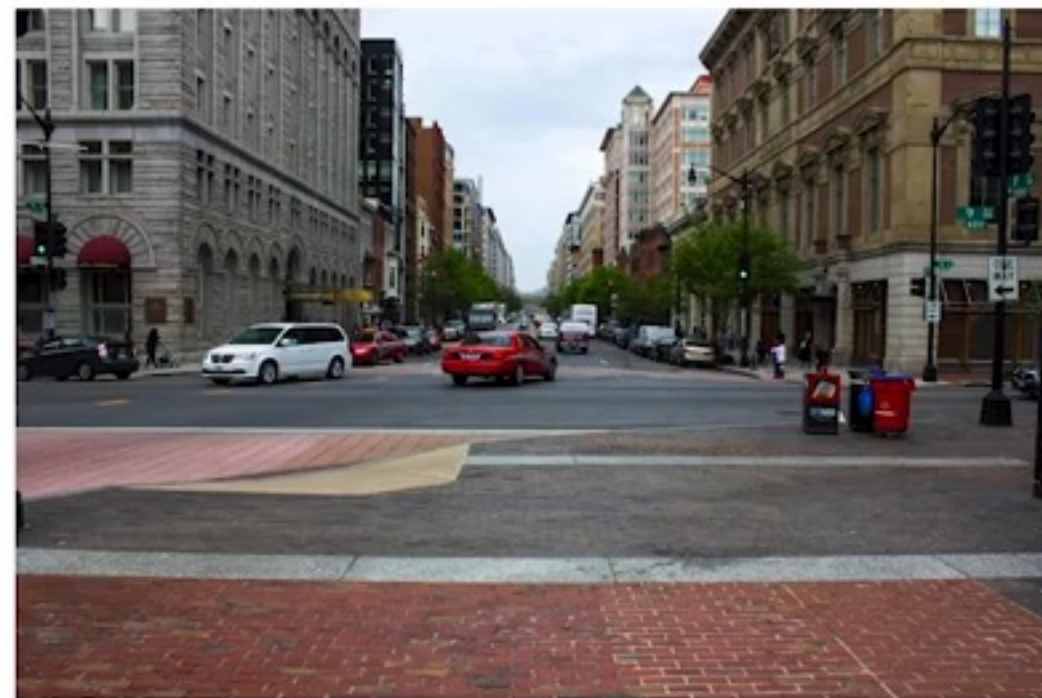
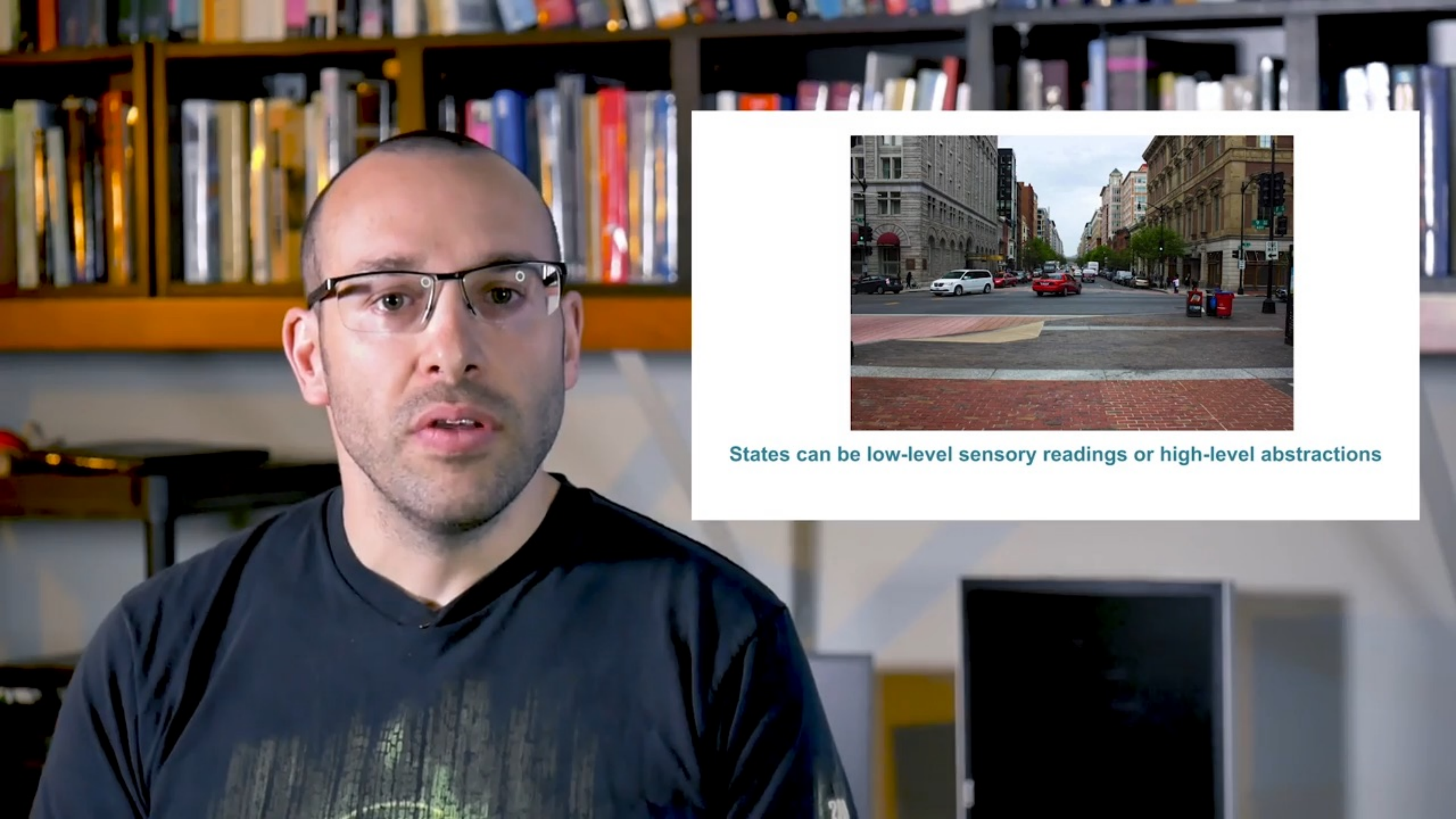
# Dynamics of the Recycling Robot

recharging the battery restores the  
energy level to high



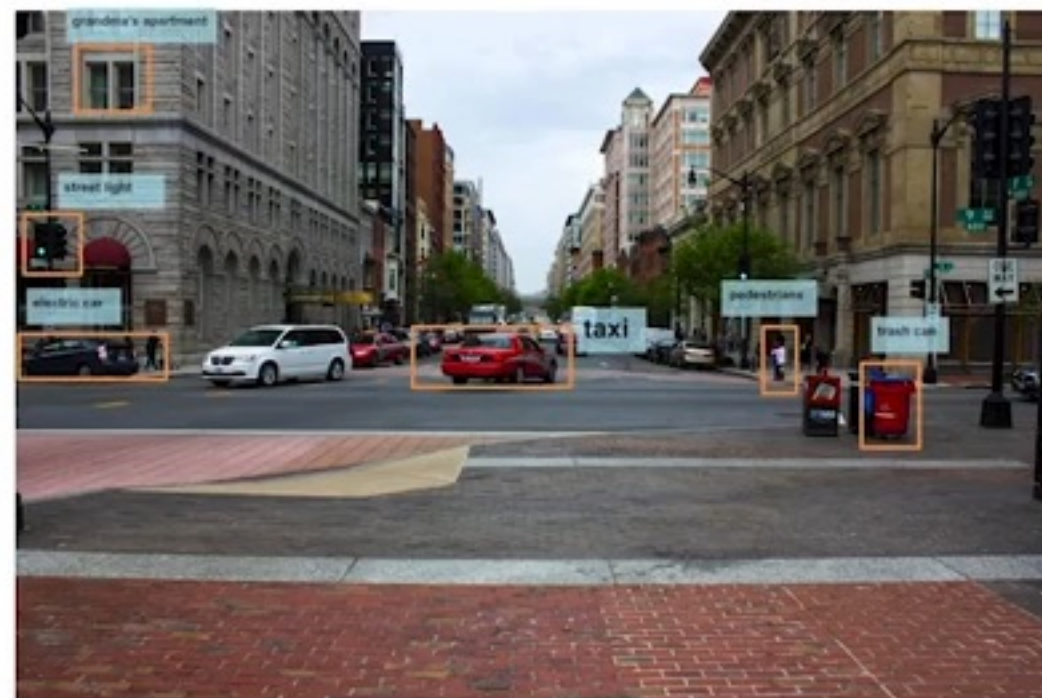
**MDP formalism is abstract and flexible**





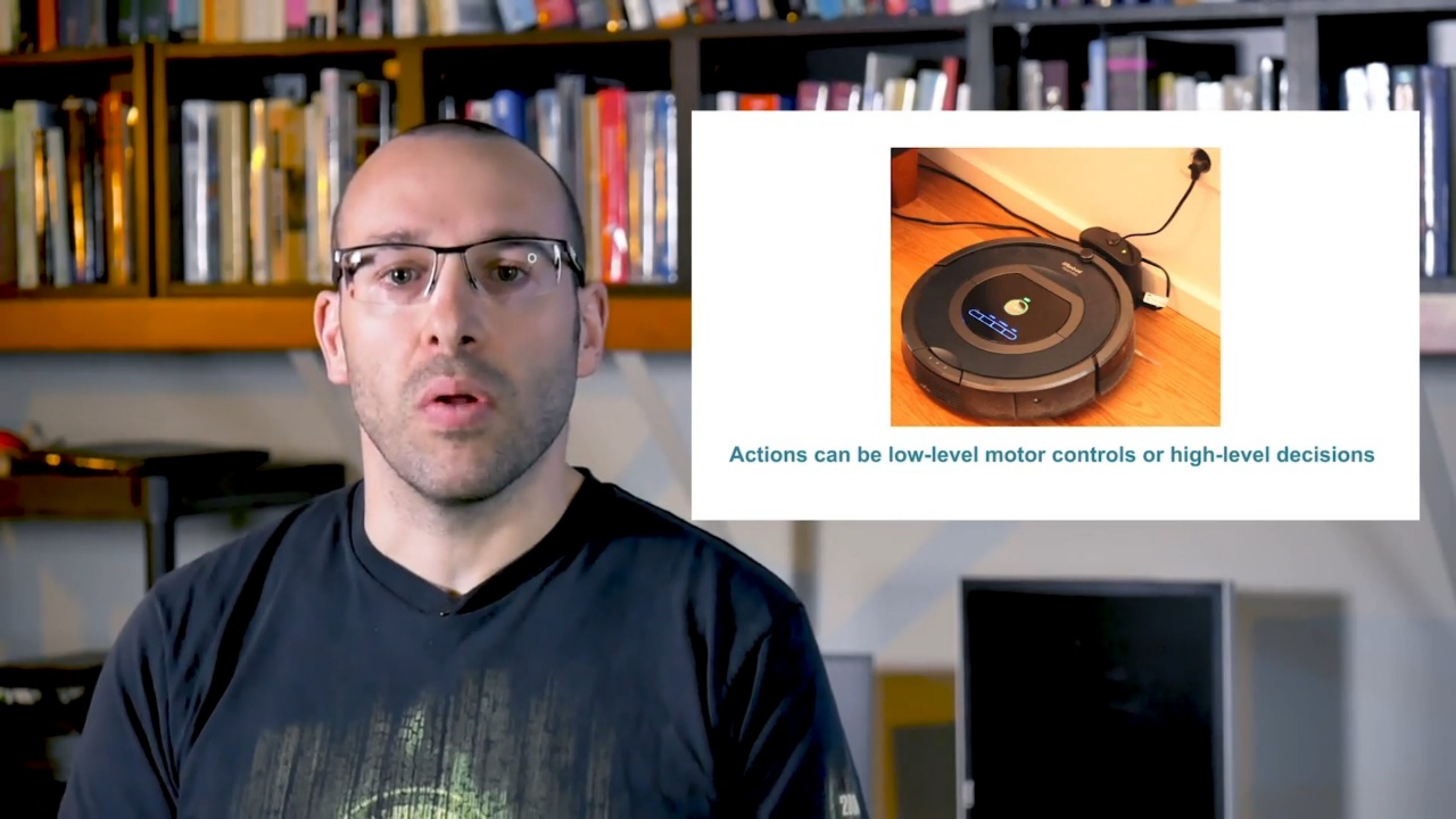
States can be low-level sensory readings or high-level abstractions





States can be low-level sensory readings or high-level abstractions





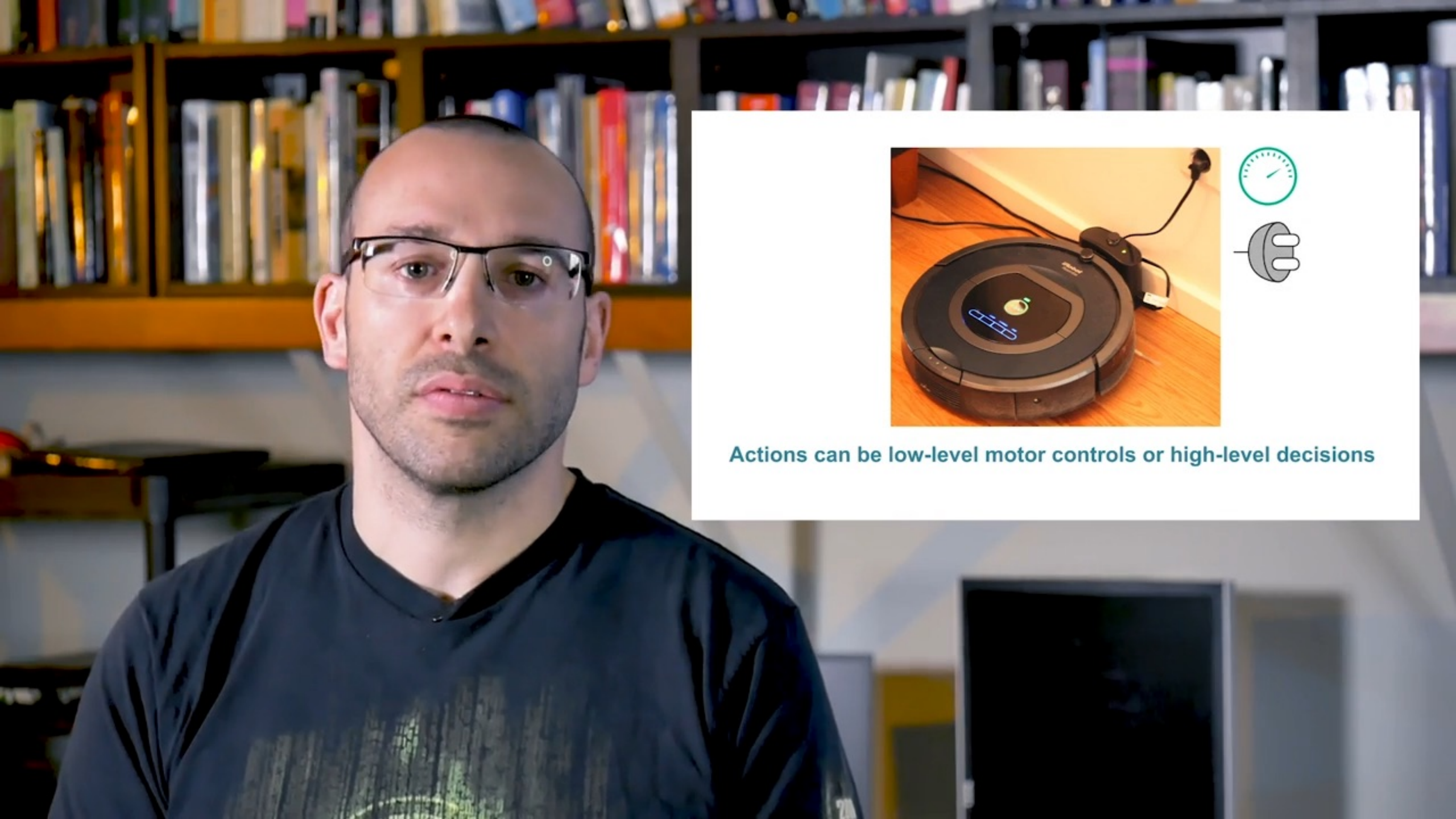
Actions can be low-level motor controls or high-level decisions





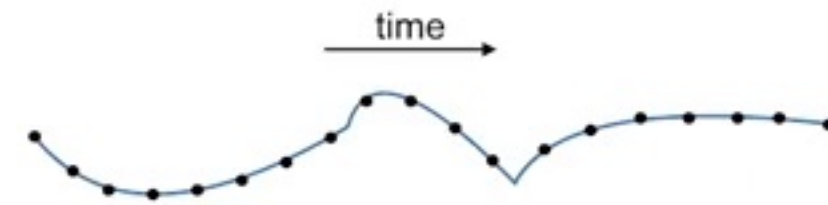
Actions can be low-level motor controls or high-level decisions





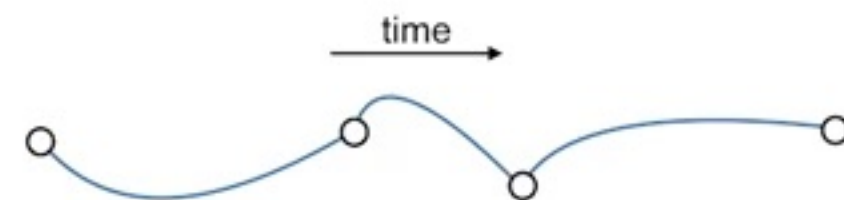
Actions can be low-level motor controls or high-level decisions





Time-steps can be fixed intervals of time  
or successive stages of decision making





**Time-steps can be fixed intervals of time  
or successive stages of decision making**





**Task:** The goal of the robot is to pick-and-place objects







**Task:** The goal of the robot is to pick-and-place objects

**State:** latest readings of joint angles and velocities







**Task:** The goal of the robot is to pick-and-place objects

**State:** latest readings of joint angles and velocities

**Action:** the amount of voltage applied to each motor







**Task:** The goal of the robot is to pick-and-place objects

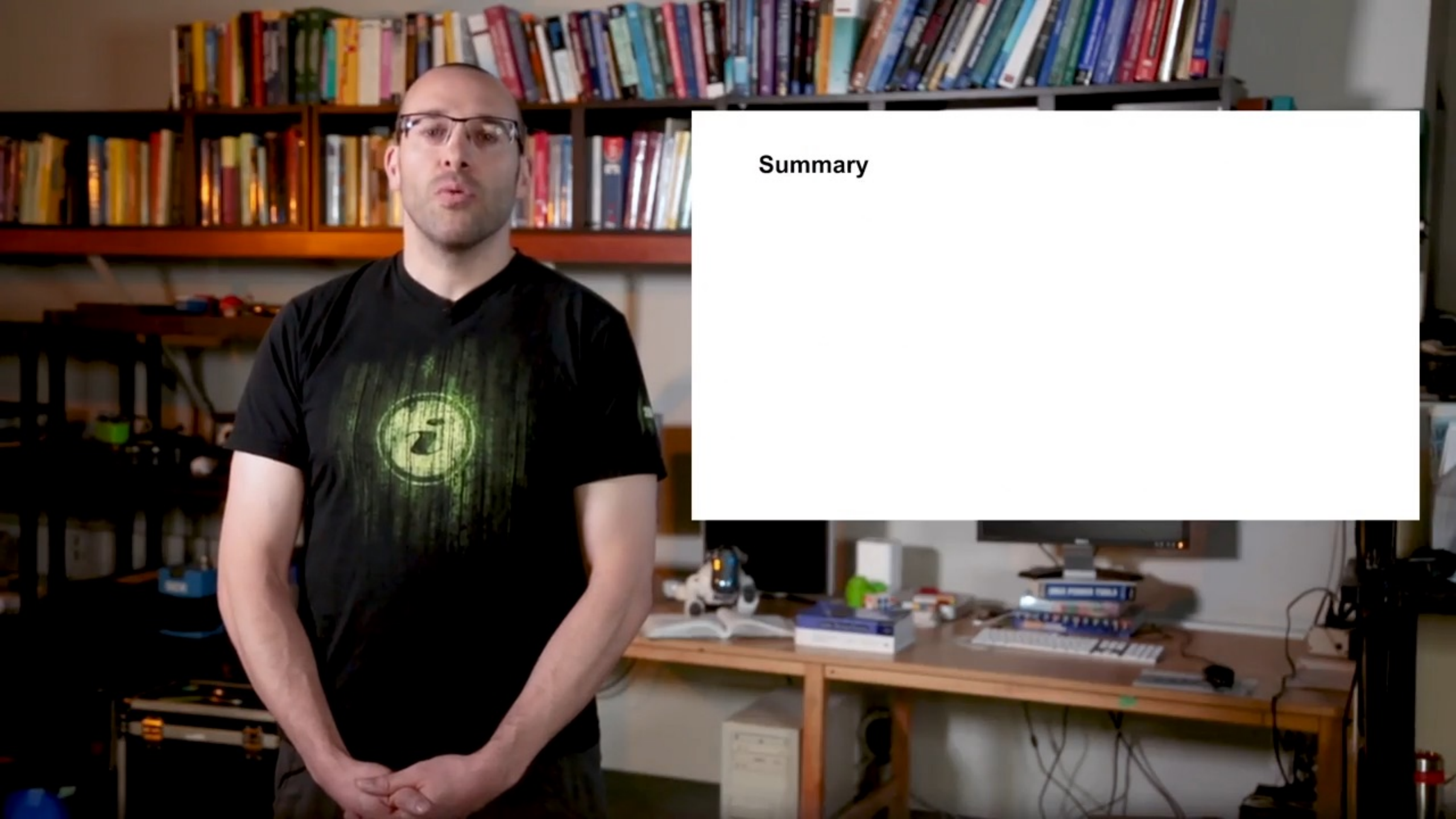
**State:** latest readings of joint angles and velocities

**Action:** the amount of voltage applied to each motor

**Reward:** +100 when an object is successfully placed  
-1 for each unit of energy consumed

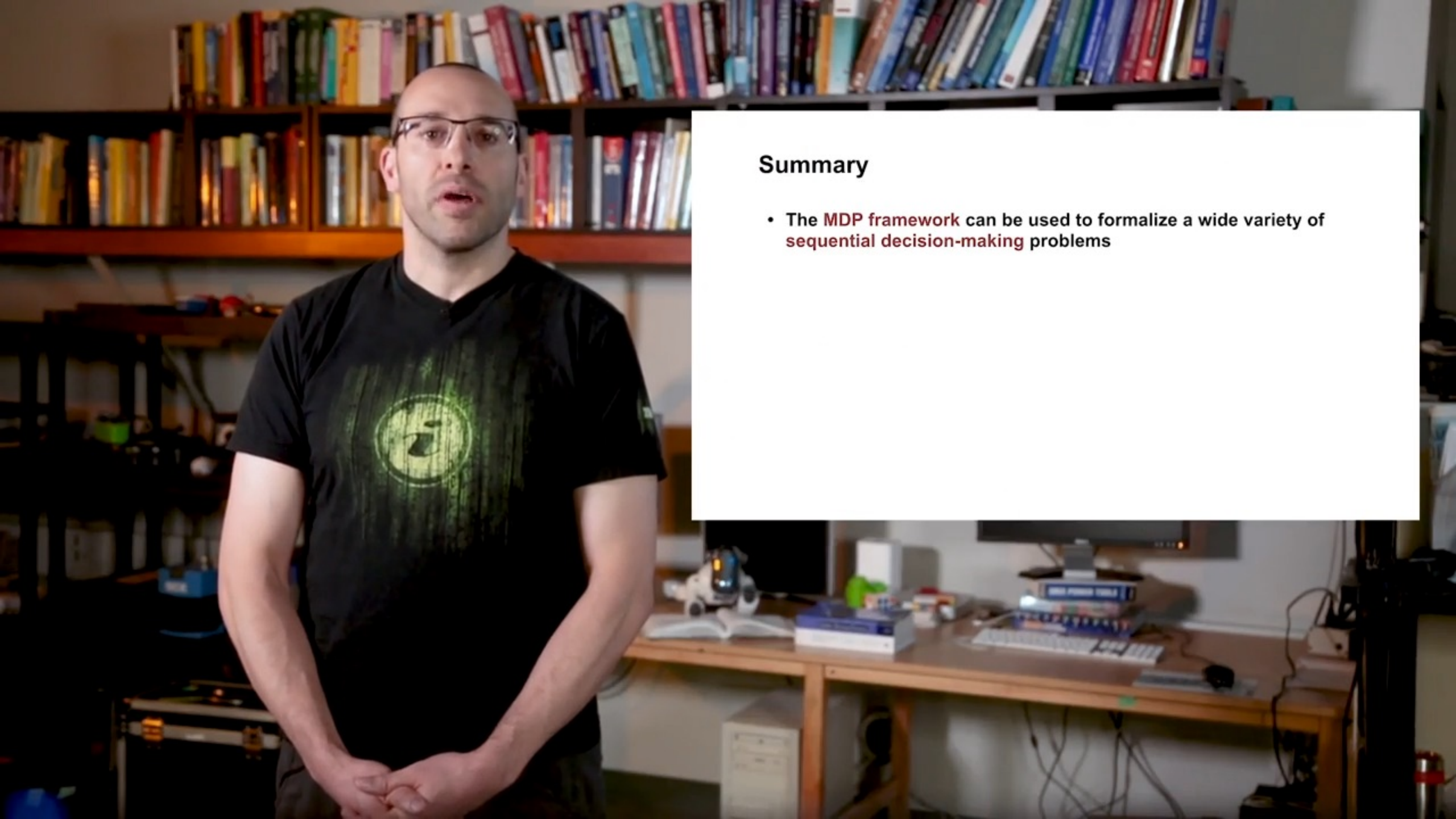






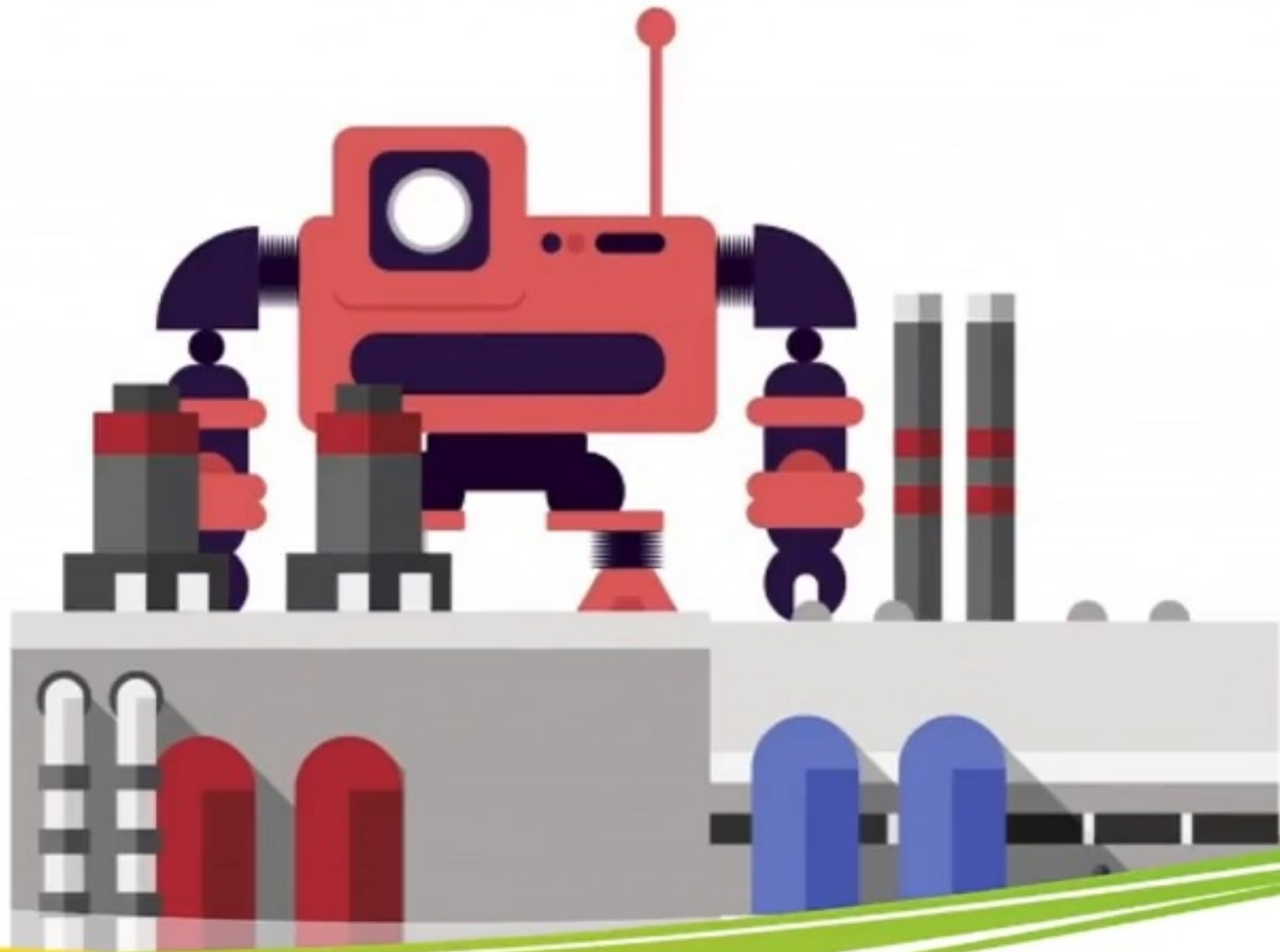
## Summary





## Summary

- The **MDP framework** can be used to formalize a wide variety of **sequential decision-making** problems



Fundamentals of Reinforcement Learning

## Examples of MDPs