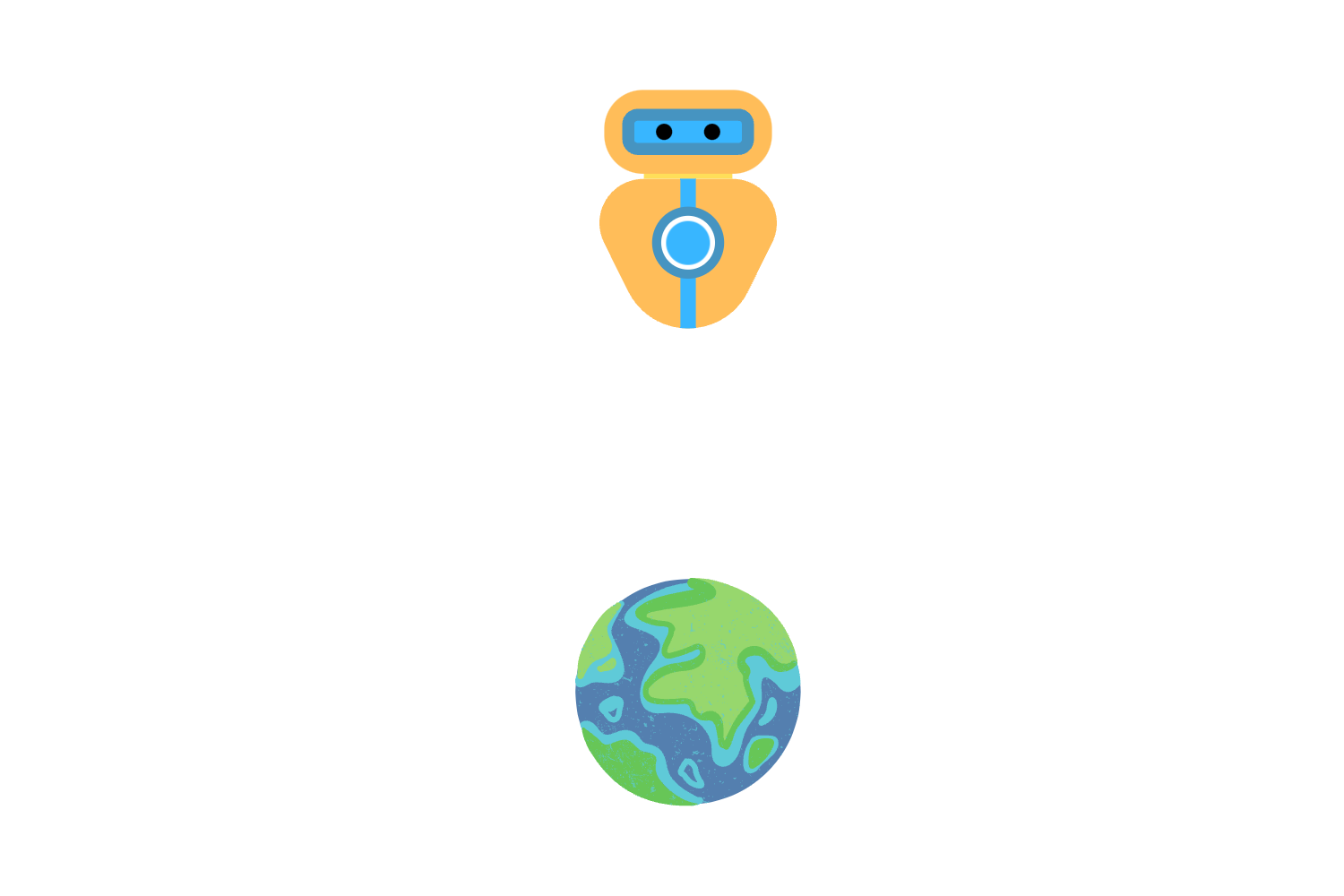
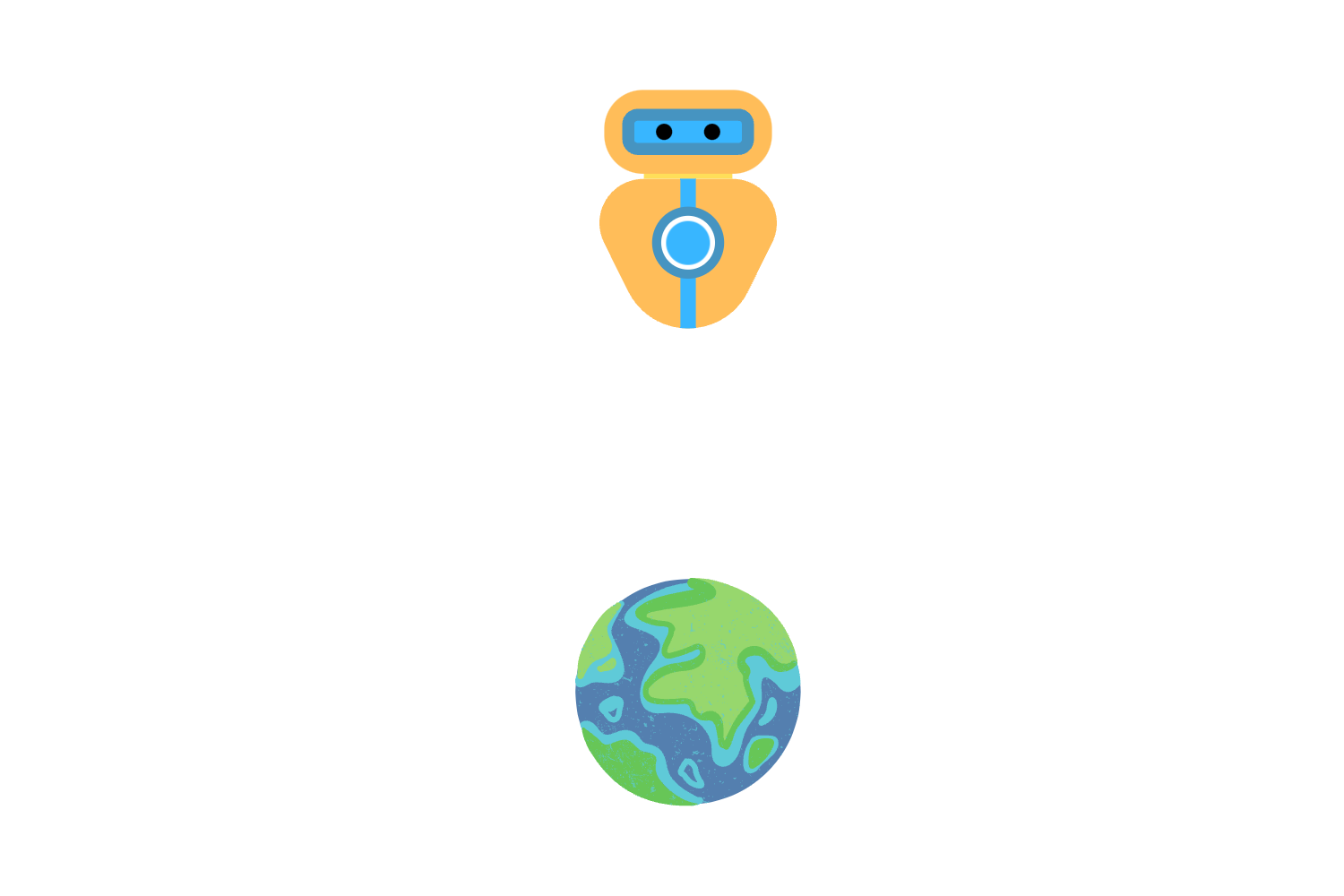
****

Agent

Observation of state

Reward

Action

****

Environment

Initialize Q-value table and hyper-parameters

Present sound to user and probe user for feedback and

Calculate uncertainty for each action

Select sound using

Terminate

Calculate change in

Calculate reward andupdate Q-value table with

Increment budget

Increment timestep

NO

YES

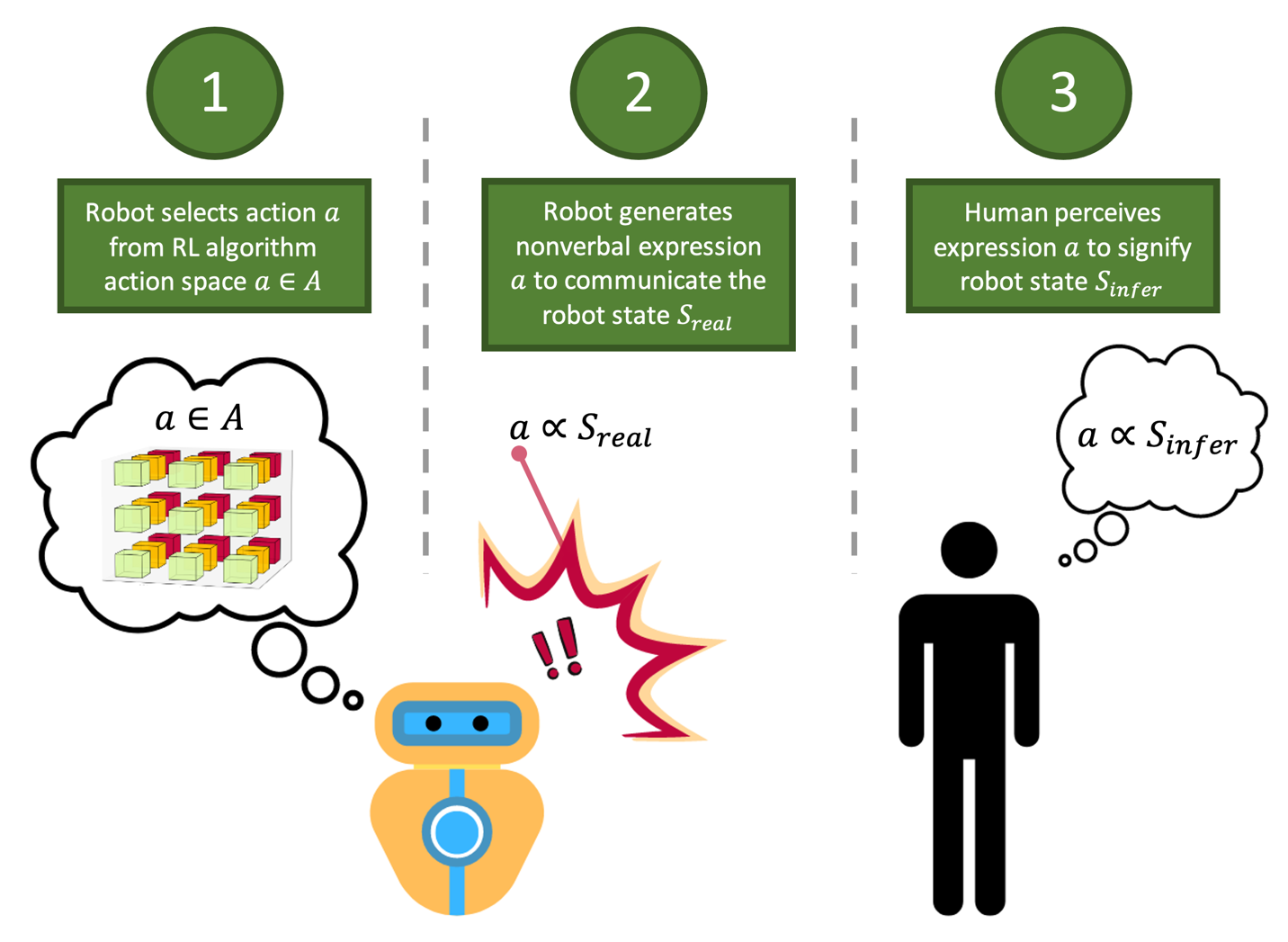
Remaining budget ?

Converged?

NO

YES

Check for convergence using

****

1

Robot selects action from RL algorithm action space

2

Robot generates nonverbal expression to communicate the robot state

3

Robot selects action from RL algorithm action space

3

Robot selects action from RL algorithm action space

**A person and a robot with speech bubbles

Description automatically generated­**