Mini-BEHAVIOR: A Procedurally Generated Benchmark for Long-Horizon Decision-Making in Embodied AI

Emily Jin*[1]   Jiaheng Hu*[2]   Zhuoyi Huang[1]   Ruohan Zhang[1]
Jiajun Wu[1]   Li Fei-Fei[1]   Roberto Martín-Martín[2]

Stanford University[1]   University of Texas, Austin[2]

RobIn
ROBOTIC INTERACTIVE
INTELLIGENCE LAB

NEURAL INFORMATION
PROCESSING SYSTEMS

Mini-BEHAVIOR is a novel fast, generative benchmark with diverse, complex tasks suitable for developing long-horizon embodied AI solutions.

Overview of Embodied AI Benchmarks

Complex Embodied AI Benchmarks
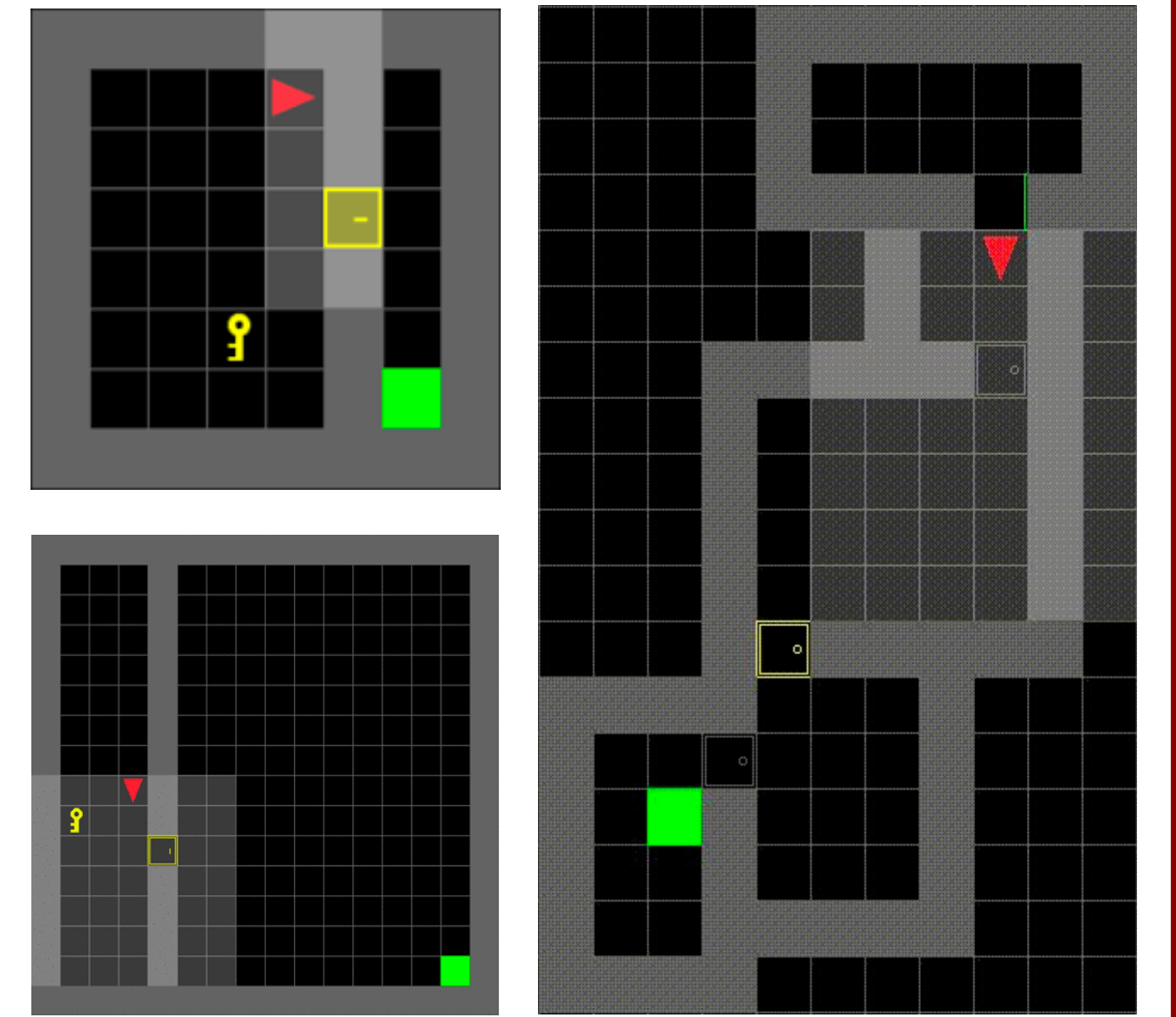
Realistic
Diverse
Complex

Reduced
speed
Huge costs

Mini BEHAVIOR

Gridworlds For Decision-Making

Lacks realism,
diversity,
complexity

Simple, fast,
easy-to-use

Features of Mini-BEHAVIOR

Diverse, Complex Tasks
a standardized set of 20 tasks that require reasoning and high-level planning skills

Procedural Generation
generate unlimited activity instances with different layouts, objects, furniture for open-ended learning

Fast Simulation Environment
a simple, lightweight, easy-to-use Gridworld environment suitable for rapid prototyping

Example Activity: Preparing a Salad

Activity Elements

Initial Conditions
OnTop(lettuce, countertop)
Inside(plate, cabinet)
...
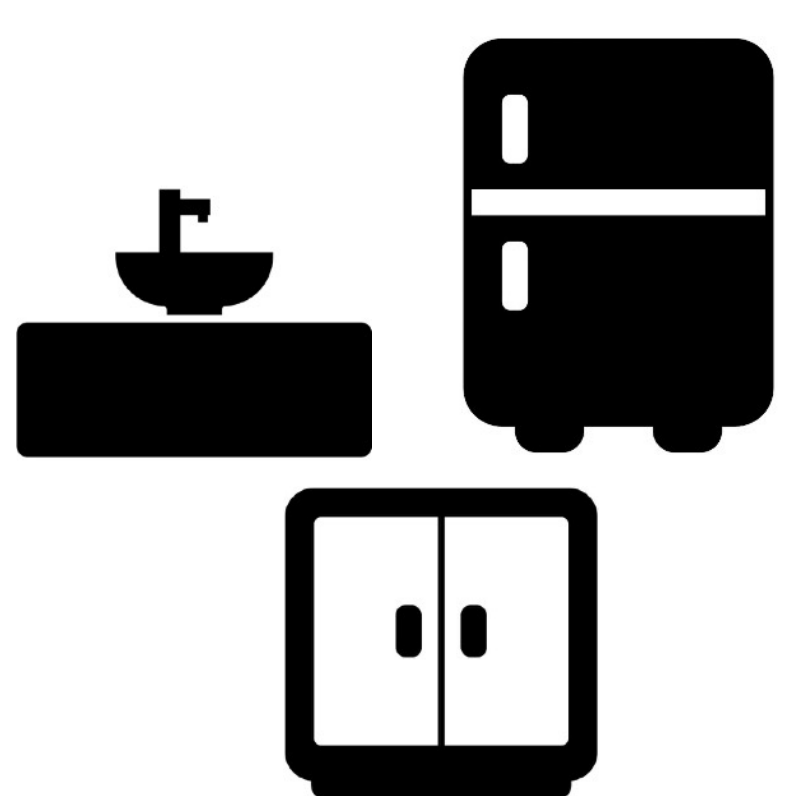**InRoom**(cabinet, kitchen)

Goal Conditions
OnTop(radish, plate)
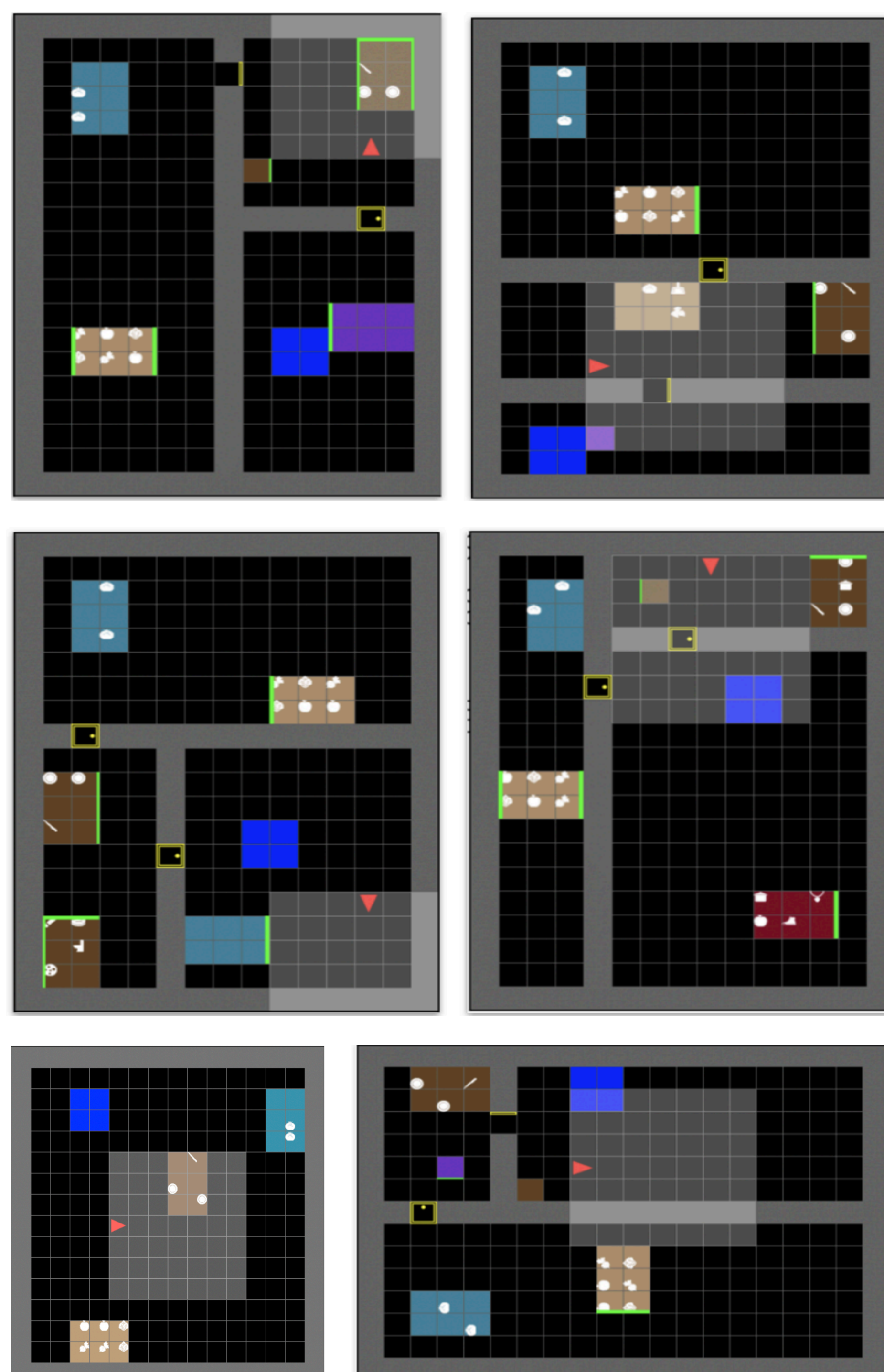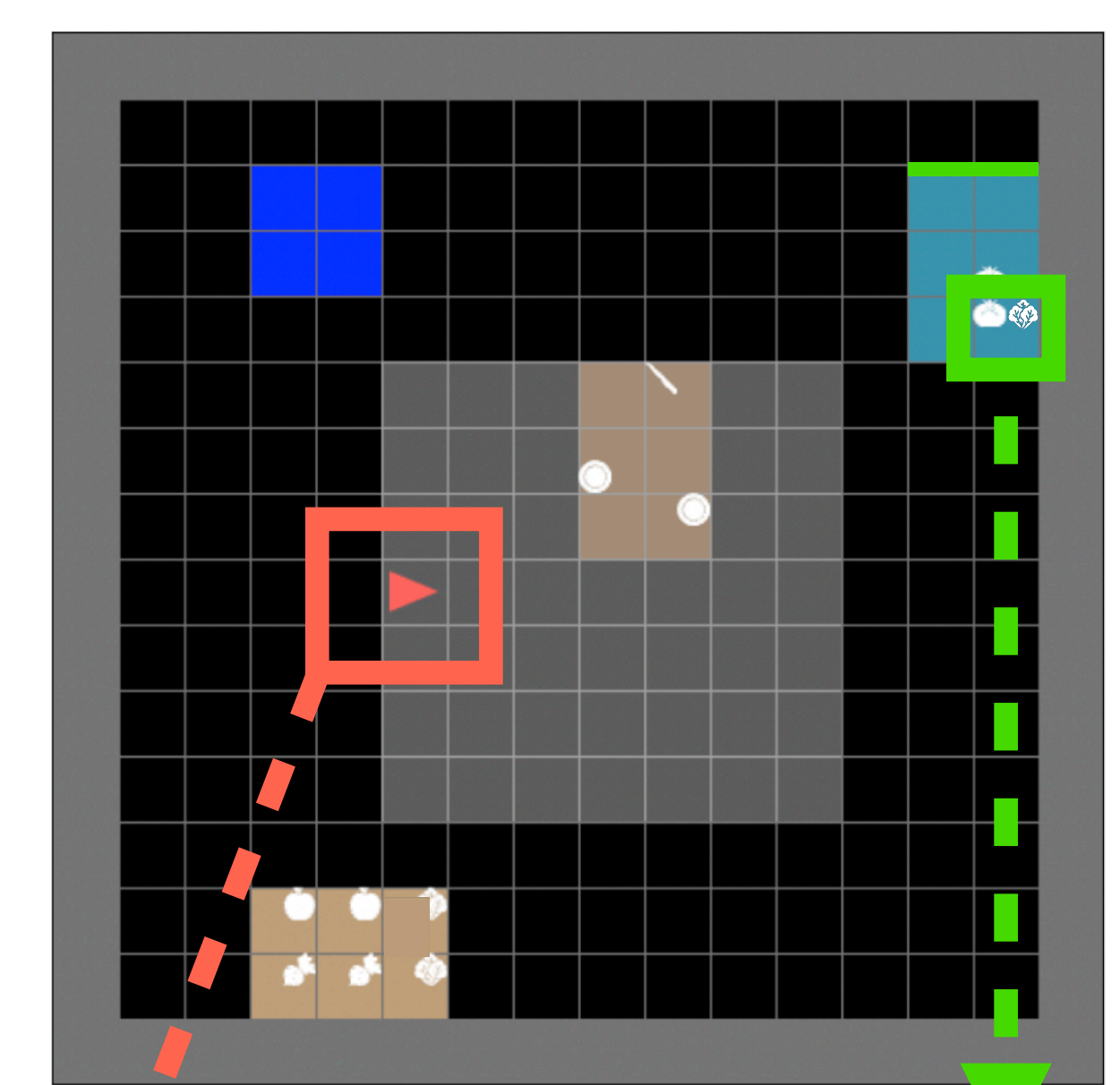**OnTop**(apple, plate)
...
Sliced(apple)

Furniture   **Objects**

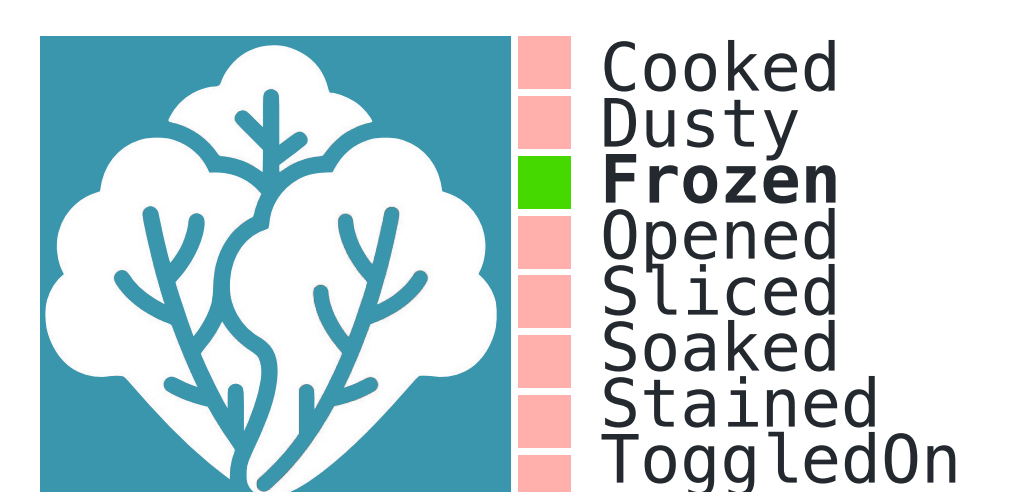### Generated Instances for Preparing a Salad

Activity Simulation Details

Agent Actions
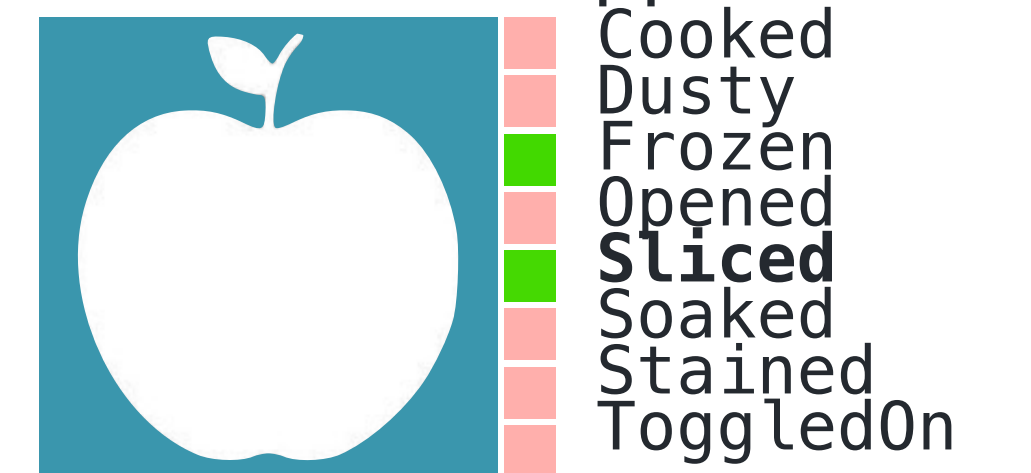
Move forward
Turn left
Turn right
Close
Cook
Drop
Drop inside
Open
Pickup
Toggle

Object/State Visuals

Frozen(lettuce)

Cooked
Dusty
Frozen
Opened
Sliced
Soaked
Stained
ToggledOn

Frozen(apple)
Sliced(apple)

Cooked
Dusty
Frozen
Opened
Sliced
Soaked
Stained
ToggledOn

While simple, Mini-BEHAVIOR is challenging for cutting-edge decision-making algorithms. We hope the embodied AI community finds value in using it to prototype and benchmark decision-making algorithms (RL, IL, HRL, etc) to ultimately solve tasks in the real world.

Code And Paper Here