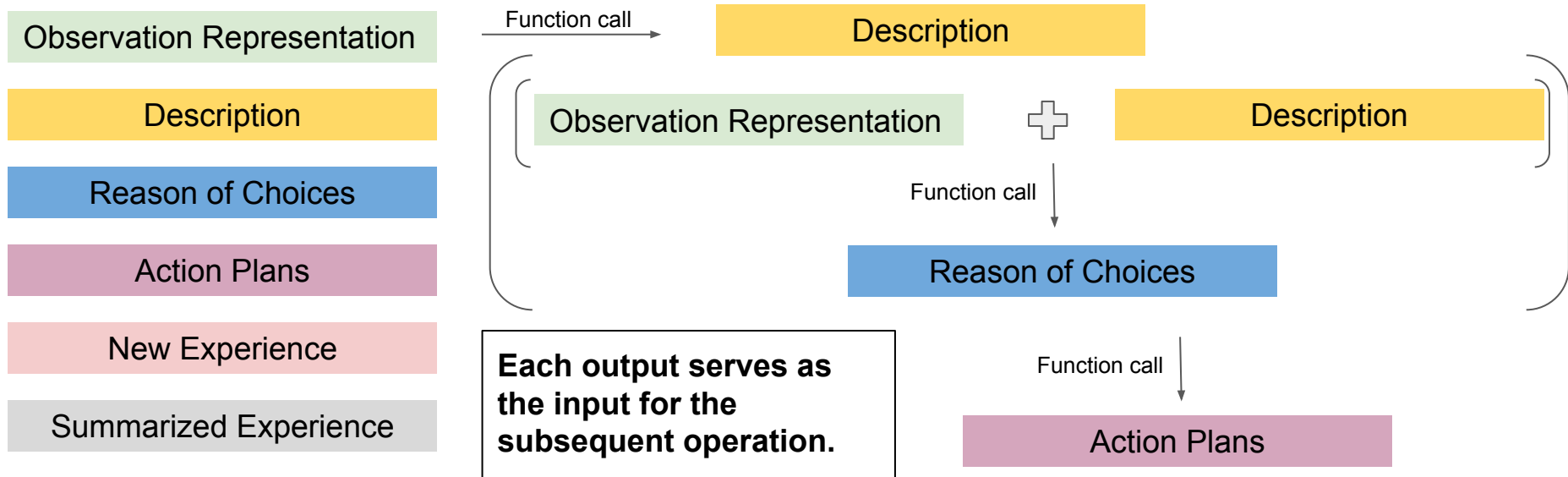


# Evaluation & Memory Architecture

# Revised Training Pipeline

**Consistency:** LLMs consistently review the entire message to provide descriptions of the environmental representation, rationale for decisions, action plans, recent experiences, and summarized learnings.



# Revised Training Pipeline

**Multiple Actions:** LLMs can now determine multiple actions based on varying situations. Consequently, a 1000-step experiment might require only half the function calls if LLMs execute one action at a time.

**Concise Expression:** We've consolidated three information channels into one, optimizing token usage without compromising information, making it more comprehensible for LLMs.

To denote object statuses, I employ symbols such as + and \* to indicate "closed" or "locked" respectively, while the absence of a symbol indicates an "open" status.

Colors are represented by numbers, **and direct object names are used in the world map to provide clear positional context.** For instance, stating "key is behind ball" is more intuitive than saying "number 5 is behind 3", followed by a lookup of their meanings, like "5 is grey" translating to "5Wall" as "grey wall".

# Training Video: We present a video demonstrating the updated training pipeline in action on MiniGrid Environment #0, spanning 1,000 steps.

## Description by agent:

You are currently in a room with grey walls. There is an empty space in front of you, and to your left is a green ball. There is also a locked green door four spaces in front of you. You are not holding anything in your

Length: 50 Limit: 50

## New Experience Obtained:

From my actions of moving forward and turning left, I have discovered that there is an empty space in front of me and to my left. There is also a green ball in the space to my left.

Length: 41 Limit: 50

## Summarized Experience:

Comparing my new experience with my previous experience, I have confirmed that the space in front of me is empty and there is a green ball to my left. This information helps me understand the layout of the world and the objects within it. By continuing to explore in different directions, I can gather more information and potentially uncover new objects or areas. This will aid in further understanding and navigation of the world.

Length: 81 Limit: 100

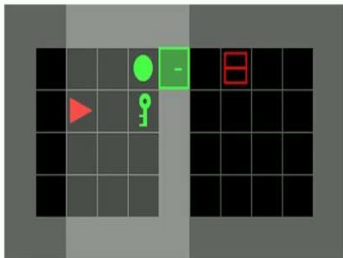
## Environment View:

```
[[0 1 2 2 2 0 0 0 0 0]
[0 1 2 2 2 0 0 0 0 0]
[0 1 2 2 2 0 0 0 0 0]
[0 1 2 2 2 0 0 0 0 0]
[0 1 2 2 2 0 0 0 0 0]
[0 1 2 2 2 0 0 0 0 0]]
```

## Environment Step:

```
[[0 0 0 0 0 0 0 0 0 0]
[0 0 0 0 0 0 0 0 0 0]
[0 1 1 0 0 0 0 0 0 0]
[0 0 0 0 0 0 0 0 0 0]
[0 0 0 0 0 0 0 0 0 0]
[0 0 0 0 0 0 0 0 0 0]]
```

## Screenshot:



## Reason of choices:

I choose to do (3) forward and (0) left actions.

Reason: By moving forward, I can explore the space in front of me and potentially discover new objects or areas. Additionally, by turning left, I can expand my field of view and gather more information about the surrounding objects and their properties. This combination of actions allows me to both move forward and gather more knowledge about the environment.

Length: 82 Limit: 100

## Status:

Position = (2, 2) Action = left Memory Length Limit = 10

Direction = Right/East Inventory: nothing

Step index = 1

Choice index = 0

## World Map (Composed):

```
[[5Wall 5Wall 5Wall 5Wall 5Wall 5Wall - - - -]
5Wall 0Empty 0Empty 0Empty 1Ball 1Door+ - - - -]
-]
5Wall 0Empty Yourself 0Empty 1Key 5Wall - - - -]
- 0Empty 0Empty 0Empty 0Empty 5Wall - - - -]
- -]
[- 0Empty 0Empty 0Empty 0Empty 5Wall - - - -]
[- 5Wall 5Wall 5Wall 5Wall 5Wall - - - -]]
```

## Past actions:

## Memory Record:

```
[[0 9 10 10 10 10 0 0 0 0 0]
[0 9 10 10 10 10 0 0 0 0 0]
[0 9 10 10 10 10 0 0 0 0 0]
[0 9 10 10 10 10 0 0 0 0 0]
[0 9 10 10 10 10 0 0 0 0 0]
[0 9 10 10 10 10 0 0 0 0 0]]
```

## World Map (Object):

```
[[2 2 2 2 2 - - - -]
2 1 1 1 6 4 - - - -]
2 1 Yourself 1 5 2 - - - -]
- 1 1 1 1 2 - - - -]
- 1 1 1 1 2 - - - -]
[- 2 2 2 2 2 - - - -]]
```

## Metrics:

Environment View % = 45.45454545454545

Environment Step % = 3.0303030303030303

Object Toggle % = 0.0

Object Drop Off % = 0.0

Object Pick Up % = 0.0

# Evaluation Methodology

We assessed the trained agent based on its last-step experience, which includes the following:

- 1. Toggling objects can reveal hidden doors or paths, leading to new areas or objects.*
- 2. Moving forward allows me to uncover unseen areas and potentially discover new objects or paths.*
- 3. Changing my facing direction provides different perspectives and helps uncover previously unseen objects or paths.*
- 4. It is important to thoroughly explore each direction to ensure I don't miss any hidden objects or paths.*

For the assessment, the descriptions remain consistent with those used during the agent's training. Thus, for **identical positions and observation representations, the description remains the same**. The evaluation only encompasses the first three stages: "Observation Representation -> Description -> Reason for Choices". We **anticipate that the agent will leverage its prior experience to achieve the provided objective**. For instance, its experience of successfully opening a door in the past might guide its current action.

Additionally, we find it pertinent to juxtapose this performance against a scenario where the LLMs are directly tasked to formulate an action plan based on the given goal. Specifically, we're interested in observing how the LLMs operate with just the environment's observation representation and **without the benefit of learned experience for a comparative analysis**.

# Evaluation Settings

This is the observation used in the evaluation:

You are at (2, 2), facing East/Right. You are holding nothing.

World Map:

[[5Wall 5Wall	5Wall	5Wall	5Wall	5Wall	5Wall	-	-	-	-	-	]
[5Wall 0Empty 0Empty 0Empty 1Ball	1Door+	-	-	-	-	-	-	-	-	-	]
[5Wall 0Empty Yourself 0Empty 1Key	5Wall	-	-	-	-	-	-	-	-	-	]
[ -	0Empty 0Empty 0Empty 0Empty 5Wall	-	-	-	-	-	-	-	-	-	]
[ -	0Empty 0Empty 0Empty 0Empty 5Wall	-	-	-	-	-	-	-	-	-	]
[ -	5Wall	5Wall	5Wall	5Wall	5Wall	-	-	-	-	-	]

# Evaluation Tasks

## Difficulty Levels for Evaluation:

### Elementary Level:

- a. Go forward
- b. Turn right
- c. Turn left
- d. Toggle object
- e. Pick up object
- f. Drop off object

### Very Easy Level:

- a. Step forward and turn right
- b. Turn left and step forward
- c. Turn right three times and move forward
- d. Turn left and then move forward one step
- e. Go forward two steps, turn right, and then turn right again

### Easy Level:

- a. Move to the front of the key
- b. Move to the front of the ball
- c. Move to the corner
- d. Move to the front of any wall
- e. Move to the bottom left corner

### Medium Level:

- a. Pick up the key
- b. Pick up the ball
- c. Pick up the key and place it elsewhere
- d. Pick up the ball and place it elsewhere
- e. Pick up the key, place it elsewhere, and then pick up the ball
- f. Pick up the ball and place it elsewhere
- g. Pick up the ball, place it elsewhere, pick up the key, and use it to open the door

### Very Hard Level:

- a. Open the door
- b. Open the door and enter it
- c. Open the door, place the key elsewhere
- d. Open the door, place the key elsewhere, and then pick up the ball
- e. Open the door, place the key elsewhere, and put the ball in the box

Note: Each level presents progressively complex tasks designed to evaluate the agent's capability and adaptability.

# Evaluation Results

Difficulty	Success Rate (%) - With Experience vs. Without Experience:	
Elementary	66.6	66.6
Very Easy	40.0	40.0
Easy	20.0	20.0
Medium	0.0	0.0
Hard	20.0	0.0
Very Hard	0.0	0.0

(Note: Values on the left represent performance with experience, and those on the right denote performance without experience.)



# Evaluation Results Interpretation

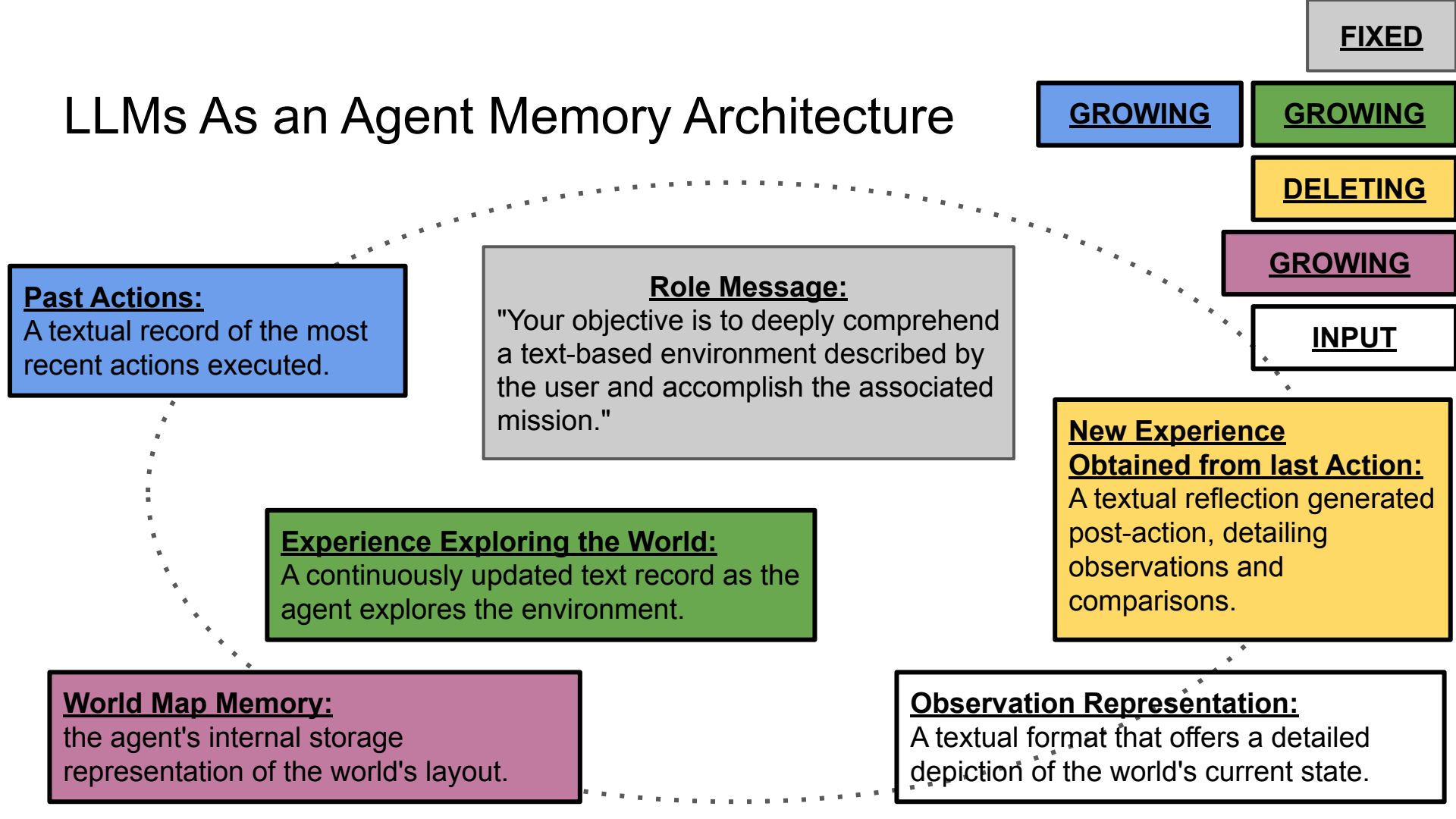
Given our one-step evaluation approach, it's understandable that the agent might not produce a flawless plan to achieve the goal in a single step.

We anticipate a significant improvement in performance if the agent is granted more steps to fulfill the objective. Furthermore, comparing the number of steps needed between agents with and without prior experience would provide a more comprehensive assessment.

It's notable that for elementary and easy goals, the agent can achieve a 100% success rate by directly utilizing functions, such as prompting the LLMs to translate "turn right three times and toggle" into an action plan, **without any reliance on observation representation knowledge**.

This highlights the significance of discerning when such knowledge is crucial for the agent's decision-making. Some tasks are inherently straightforward; for instance, when tasked with "toggle an object", an agent with observational understanding would first search for an object to toggle. In contrast, the LLMs might directly opt for the "toggle" action, favoring a more straightforward approach.

# LLMs As an Agent Memory Architecture



# LLMs As an Agent Memory Architecture

FIXED

GROWING

GROWING

DELETING

GROWING

INPUT

## Past Actions:

A textual record of the most recent actions executed.

## Role Message:

"Your objective is to deeply comprehend a text-based environment described by the user and accomplish the associated

(Note: The structure organizes the various components of the agent's memory and decision-making framework, ensuring a cohesive approach to understanding and interacting with the environment.)

A continuously updated text record as the agent explores the environment.

observations and comparisons.

## World Map Memory:

the agent's internal storage representation of the world's layout.

## Observation Representation:

A textual format that offers a detailed depiction of the world's current state.

# Limitations

Owing to the substantial time invested in refining the training pipeline for enhanced performance and reduced token consumption, certain aspects of our evaluation and architecture remain underdeveloped:

1. **Diverse Evaluations:** We could further broaden our analysis by incorporating experiments across different seeds, employing models beyond just GPT-3.5-turbo, and testing in a wider variety of environments.
2. **Memory Architecture Preliminary Stage:** Currently, our memory architecture closely resembles variable naming conventions. It hasn't matured into a holistic system yet. We encourage viewing it as a preliminary glimpse into the composition of messages used in this "LLMs as an Agent" initiative.

# Next Steps

## **1. Clarify the Project's Contribution:**

As I currently understand it, the utilization of LLMs to direct an agent's actions in task execution is a burgeoning area of interest. Remarkably, enabling LLMs to explore an environment from scratch without pre-specified tasks proves beneficial, particularly in terms of conserving training time and saving training samples.

This project delves into an innovative memory architecture for LLMs, enabling efficient environmental exploration and adaptive evolution. Our objective is to underscore the potential advantages of this architecture in enhancing the application of LLMs in robotic control.

## **2. Detail the Memory Architecture:**

It's essential to provide a comprehensive overview of the memory architecture being employed. Addressing potential limitations and proposing future avenues of exploration and refinement will be pivotal.

## **3. Experimental Validation:**

Undertake a series of experiments to validate and showcase the superiority and efficacy of the proposed memory architecture. By addressing these steps, we aim to solidify the foundational aspects of our project, setting the stage for future advancements and refinements.

# Next Steps

## 4. Beyond Textual Representations

Until now, our primary mode of representation for the agent's memory has been text-based: encompassing experience, world map, and past actions. This choice was largely influenced by the need to convey information in a format understandable by LLMs.

However, a pertinent question arises: Can LLMs interpret more comprehensive, multi-modal data? We should consider exploring a fusion of prior actions, observations, and experiences and investigate the feasibility of directly saving and loading such multi-modal information into the LLMs model.

If we can integrate diverse forms of information, such as visual data or neural network weights, alongside textual data, it could provide a more efficient method for LLMs to make decisions and achieve a deeper understanding.

# Some Evaluation Result (with experience)

Elementary Goals	Action	Reason	Success
Go forward	Forward	I choose to do (3) forward to move myself to the front object if it's empty or opened. Reason: Moving forward allows me to explore and uncover unseen areas. It is a simple action that can potentially reveal new objects or paths in the environment.	True
Turn right	Right	I choose to do (1) right: turn my view to the right object. Reason: By turning my view to the right, I can gain a different perspective of the surrounding objects and potentially uncover new paths or objects that were previously unseen. This will help me to explore the world more thoroughly and gather more information for my mission.	True

## Some Evaluation Result (with experience)

Elementary Goals	Action	Reason	Success
Turn left	Left	I choose to do (0) left: turn my view to the left object. Reason: By turning left, I can get a different perspective and potentially uncover new objects or paths that were previously unseen. This action helps me in exploring my surroundings more thoroughly and increases the chances of finding useful items or discovering hidden areas.	True
Toggle Object	Toggle	I choose to do (2) toggle: toggle the object in front of me. Reason: Toggling objects can reveal hidden doors or paths. By toggling the object in front of me, I may uncover a hidden path or door, potentially leading to new areas or objects.	True



# Some Evaluation Result (with experience)

Elementary Goals	Action	Reason	Success
Pick Up Object	Forward, Pick Up	I choose to do (3) forward and (4) pick up actions. Reason: By moving forward, I can reach the object in front of me and potentially uncover new areas or objects. Once I reach the object, I can pick it up to add it to my inventory. This combination of actions allows me to explore the world and acquire new items for future use.	True
Drop Off Object	Toggle	I choose to do (2) toggle: toggle the object in front of me. Reason: Toggling objects can reveal hidden doors or paths, which may lead to new areas or objects. This action allows me to potentially uncover unseen areas and discover new possibilities in the environment.	False

## Some Evaluation Result (with experience)

Very Easy Goals	Action	Reason	Success
Step forward and turn right	Right	I choose to do (1) right: turn my view to the right object. Reason: By turning my view to the right, I can gather more information about the objects and potential paths in that direction. This will help me make a more informed decision about my next action and potentially uncover new objects or paths to explore.	False
Turn left and step forward	Left	I choose to do (0) left: turn my view to the left object. By turning left, I can have a better view of my surroundings and potentially discover any hidden objects or paths that were previously unseen.	False

# Some Evaluation Result (with experience)

Very Easy Goals	Action	Reason	Success
Turn right three times and forward	Right, Right, Right, Forward	I choose to do the following actions: (a) right: turn my view to the right object (b) right: turn my view to the right object (c) right: turn my view to the right object (d) forward: move myself to the front object if it's empty or opened Reason: By turning right three times, I will be facing the same direction as before, but I will have explored the objects on my right side. This will help me uncover any hidden objects or	True
Turn left and then forward one step	Left	I choose to do (0) left: turn my view to the left object. By turning left, I can change my facing direction to face the green ball (1Ball) and potentially uncover any hidden objects or paths in that direction. This action allows me to gather more information about my surroundings and make informed decisions for future exploration.	False

# Some Evaluation Result (with experience)

Easy Goals	Action	Reason	Success
Go forward two steps and turn right, and turn right again	Forward, Right, Right	I choose to do the following actions: (a) Action 3 - forward (to move two steps forward) (b) Action 1 - right (to turn right) (c) Action 1 - right (to turn right again) Reason: By moving forward two steps, I can explore unseen areas and potentially discover new objects or paths. Turning right twice will help me change my direction to face the right object and continue exploring the surroundings effectively.	True
Move to the front of key	Forward	I choose to do (3) forward: move myself to the front object if it's empty or opened. Reason: By moving forward, I can approach the key and potentially reach it. This action allows me to physically move closer to the desired object in the shortest possible way.	True

Easy Goals	Action	Reason	Success
Move to the front of Ball	Forward	I choose to do (3) forward: move myself to the front object if it's empty or opened. Reason: Moving forward is the most direct action to reach the front of the ball. By moving forward, I can uncover unseen areas and potentially discover new objects or paths that may lead me closer to the ball.	False
Move to the corner	Forward	I choose to do (3) forward to move myself to the front object if it's empty or opened. Reason: Moving forward allows me to explore and uncover unseen areas of the world. By continuously moving forward, I can progress towards the corner of the world and potentially discover new objects or paths along the way.	False
Move to the front of any wall	Forward	I choose to do (3) forward action to move myself to the front object if it's empty or opened. This action allows me to explore and move towards the wall in front of me, which is my mission.	False