R2C: Mapping Room to Chessboard to Unlock LLM As Low-Level Action Planner

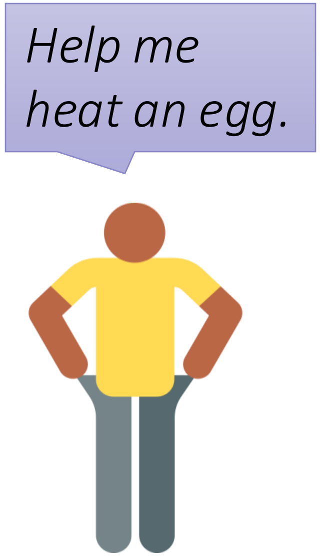
Ziyi Bai, Hanxuan Li, Bin Fu, Chuyan Xiong, Ruiping Wang, Xilin Chen

Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China

{ziyi.bai, hanxuan.li, bin.fu, chuyan.xiong}@vipl.ict.ac.cn, {wangruiping, xlchen}@ict.ac.cn

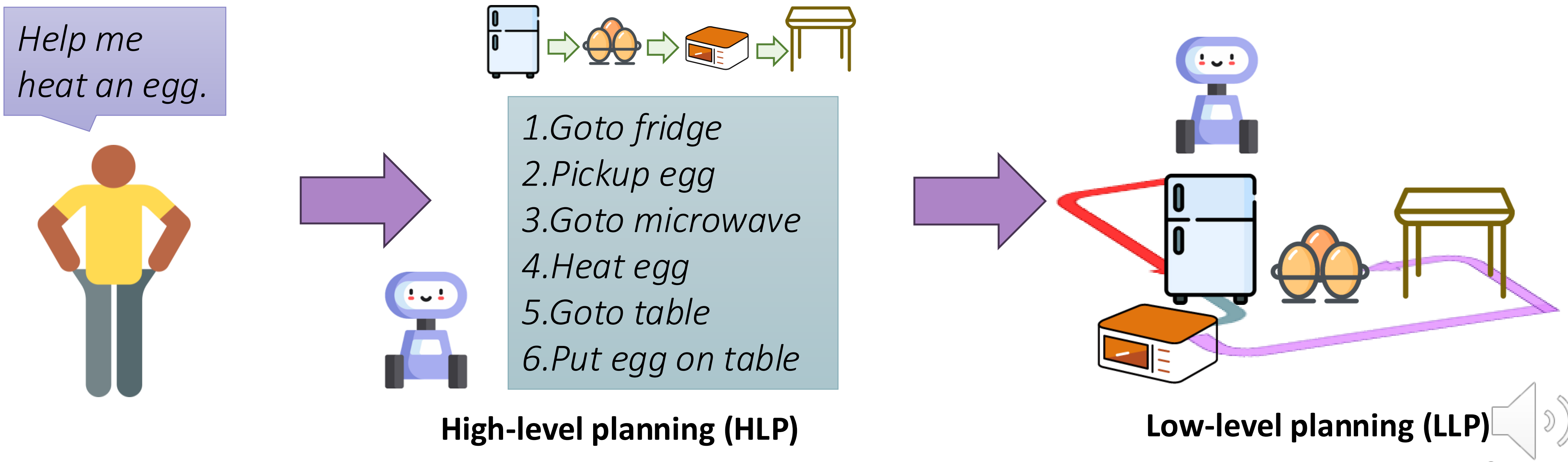
MOTIVATION

- Long-horizon Robotic Task Planning



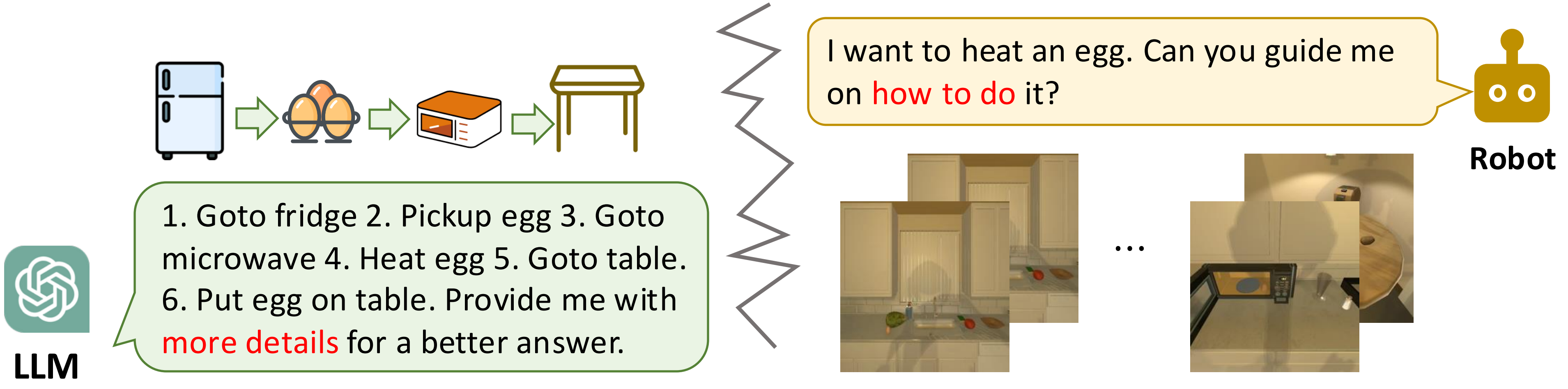
MOTIVATION

- Long-horizon Robotic Task Planning
  - Need hierarchical planning
  - High-level planning: language instruction → subgoal sequences
  - Low-level planning: subgoal → action sequences



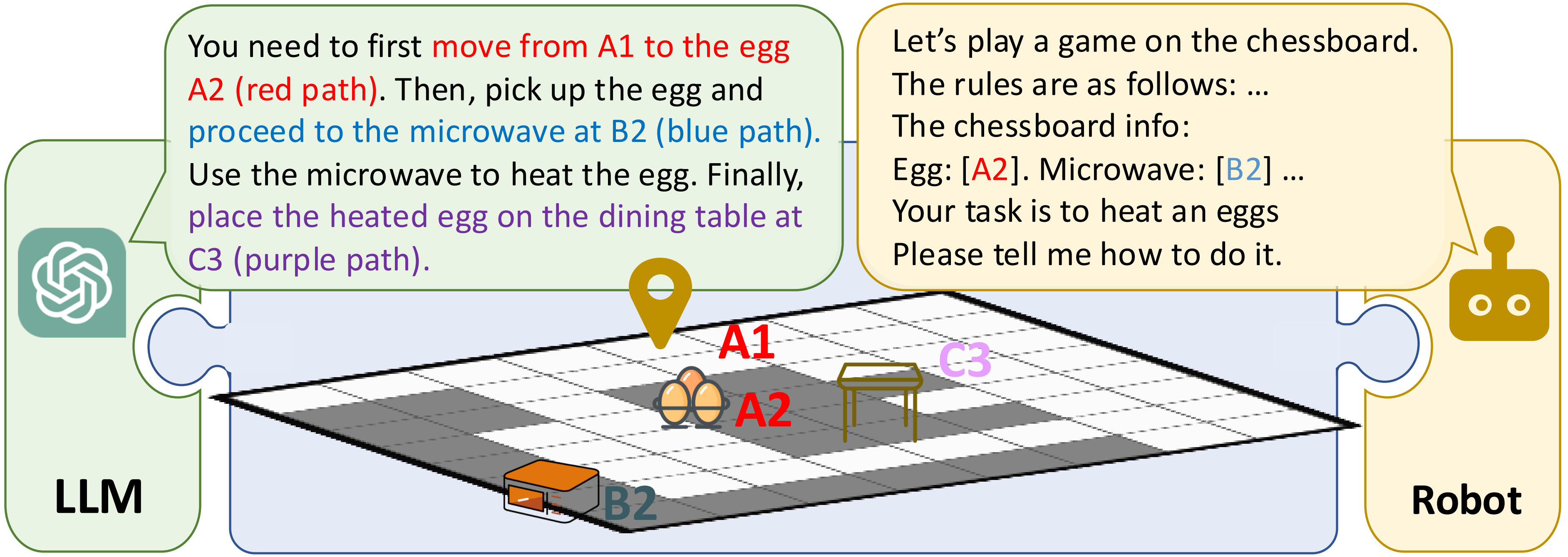
MOTIVATION

- LLMs Are Good High-Level Planners (Robot Brain)
  - Strong generalization among various tasks
  - Good world knowledge
- BUT LLMs Have No Proprioception
  - Cannot sense detailed spatial state
  - Cannot give low-level action planning

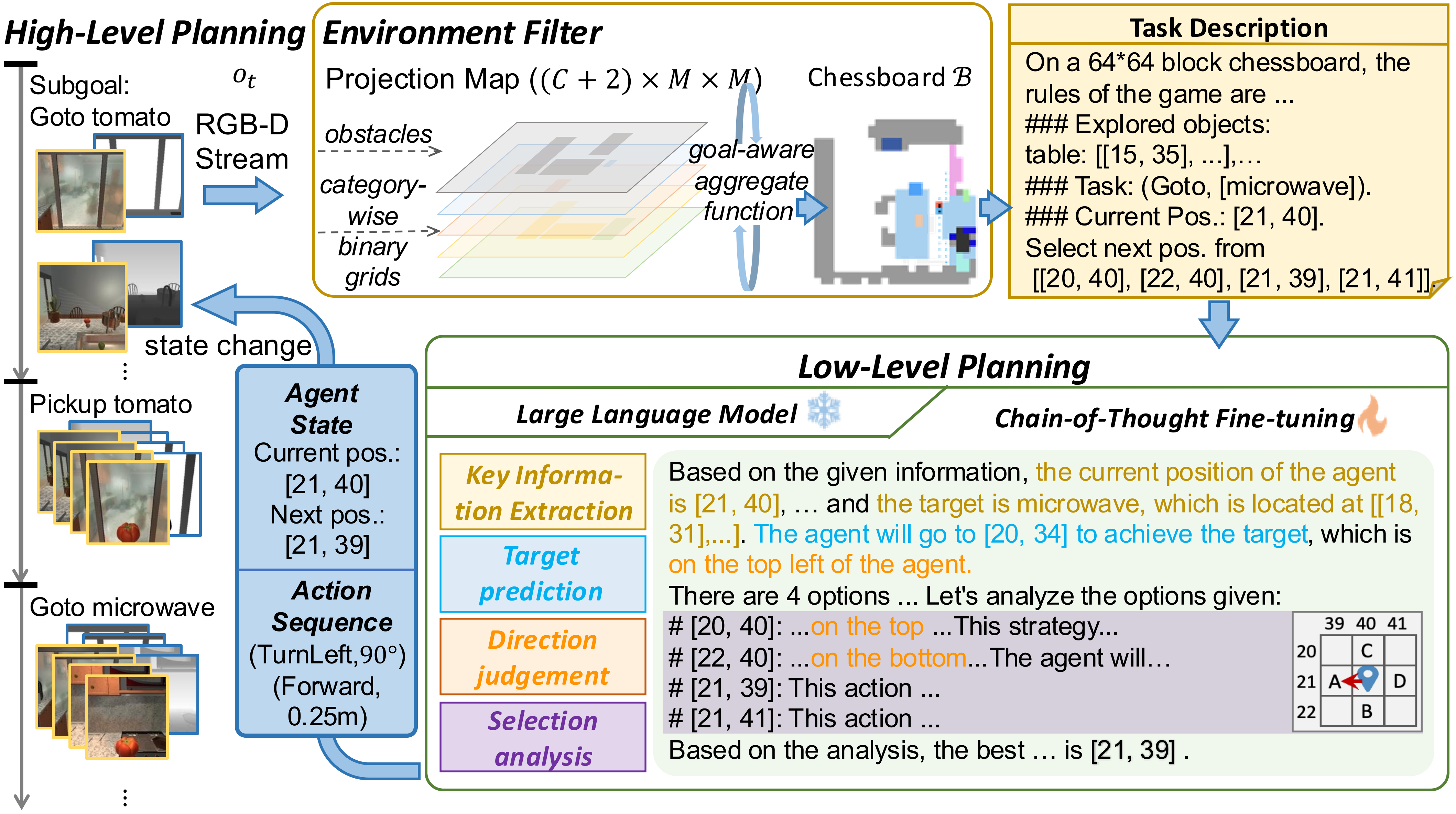


MOTIVATION

- Mapping Room to Chessboard
  - Chessboard, a common language for both LLMs and Robots
    - Robots map observations on chessboard, with its coordinates clearly show room layout
    - LLMs give low-level action plans, each cell corresponds to one step of the agent



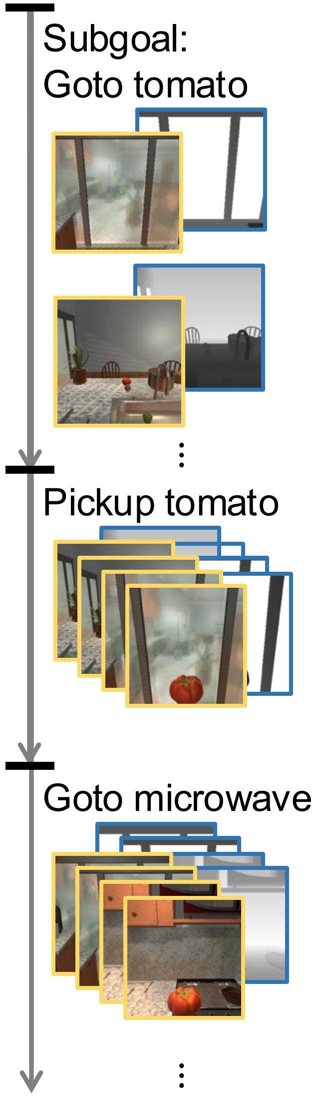
FRAMEWORK



High-Level Planning
Subgoal:
Goto tomato
o_t
RGB-D
Stream
state change
Pickup tomato
Goto microwave
Environment Filter
Projection Map $((C+2) \times M \times M)$
Chessboard $\mathcal{B}$
obstacles
category-wise binary grids
goal-aware aggregate function
Task Description
On a 64*64 block chessboard, the rules of the game are ...
Explored objects:
table: [[15, 35], ...],…
Task: (Goto, [microwave]).
Current Pos.: [21, 40].
Select next pos. from
[[20, 40], [22, 40], [21, 39], [21, 41]].
Agent State
Current pos.: [21, 40]
Next pos.: [21, 39]
Action Sequence
(TurnLeft,90°)
(Forward, 0.25m)
Low-Level Planning
Large Language Model
Chain-of-Thought Fine-tuning
Key Information Extraction
Target prediction
Direction judgement
Selection analysis
Based on the given information, the current position of the agent is [21, 40], … and the target is microwave, which is located at [[18, 31],...]. The agent will go to [20, 34] to achieve the target, which is on the top left of the agent.
There are 4 options ... Let's analyze the options given:
[20, 40]: ...on the top ...This strategy...
[22, 40]: ...on the bottom...The agent will…
[21, 39]: This action ...
[21, 41]: This action ...
Based on the analysis, the best … is [21, 39] .
39 40 41
20 C
21 A D
22 B

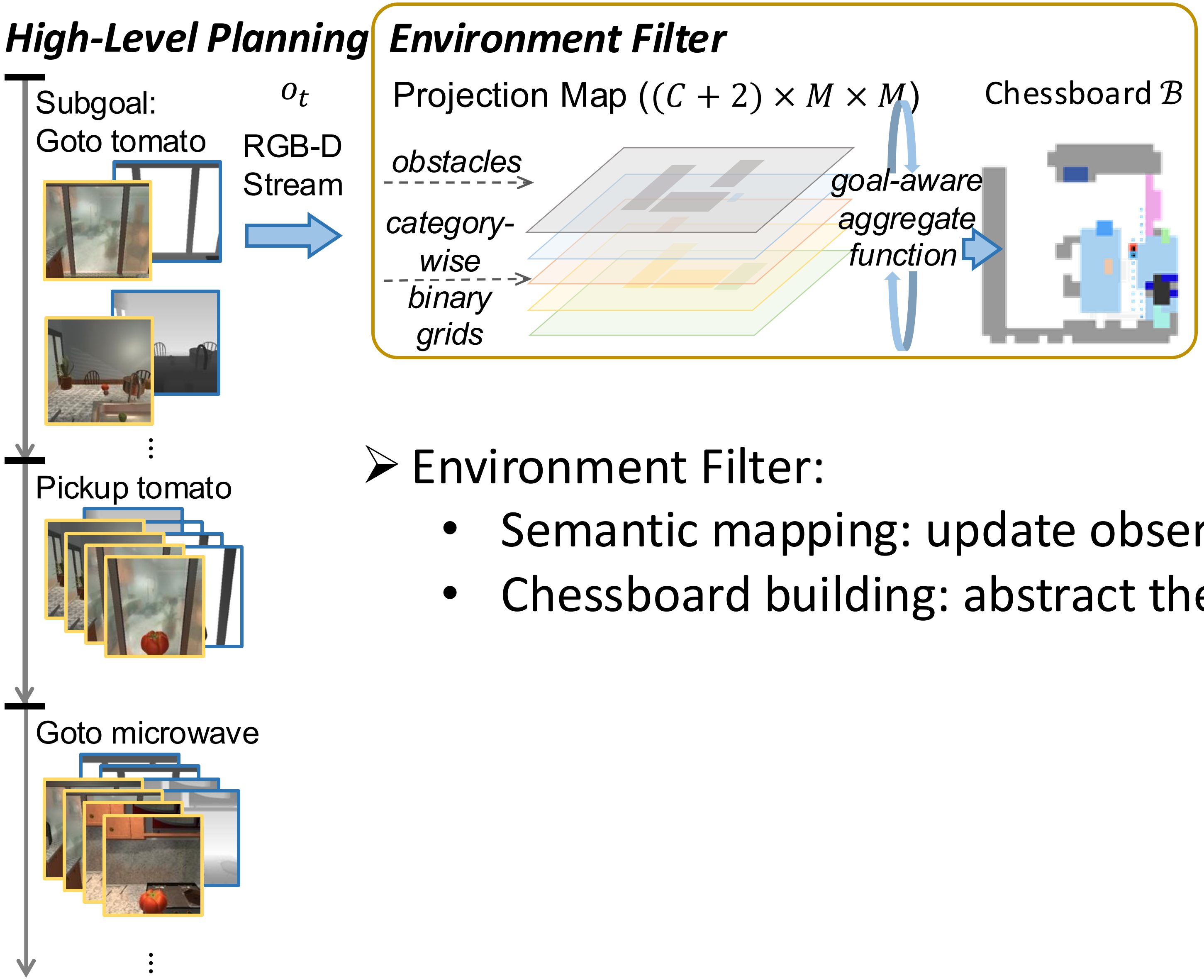
FRAMEWORK

High-Level Planning



- High-Level Planning (HLP): decompose long-horizon task into a sequence of subgoals G= [G_1,G_2,…, G_K]

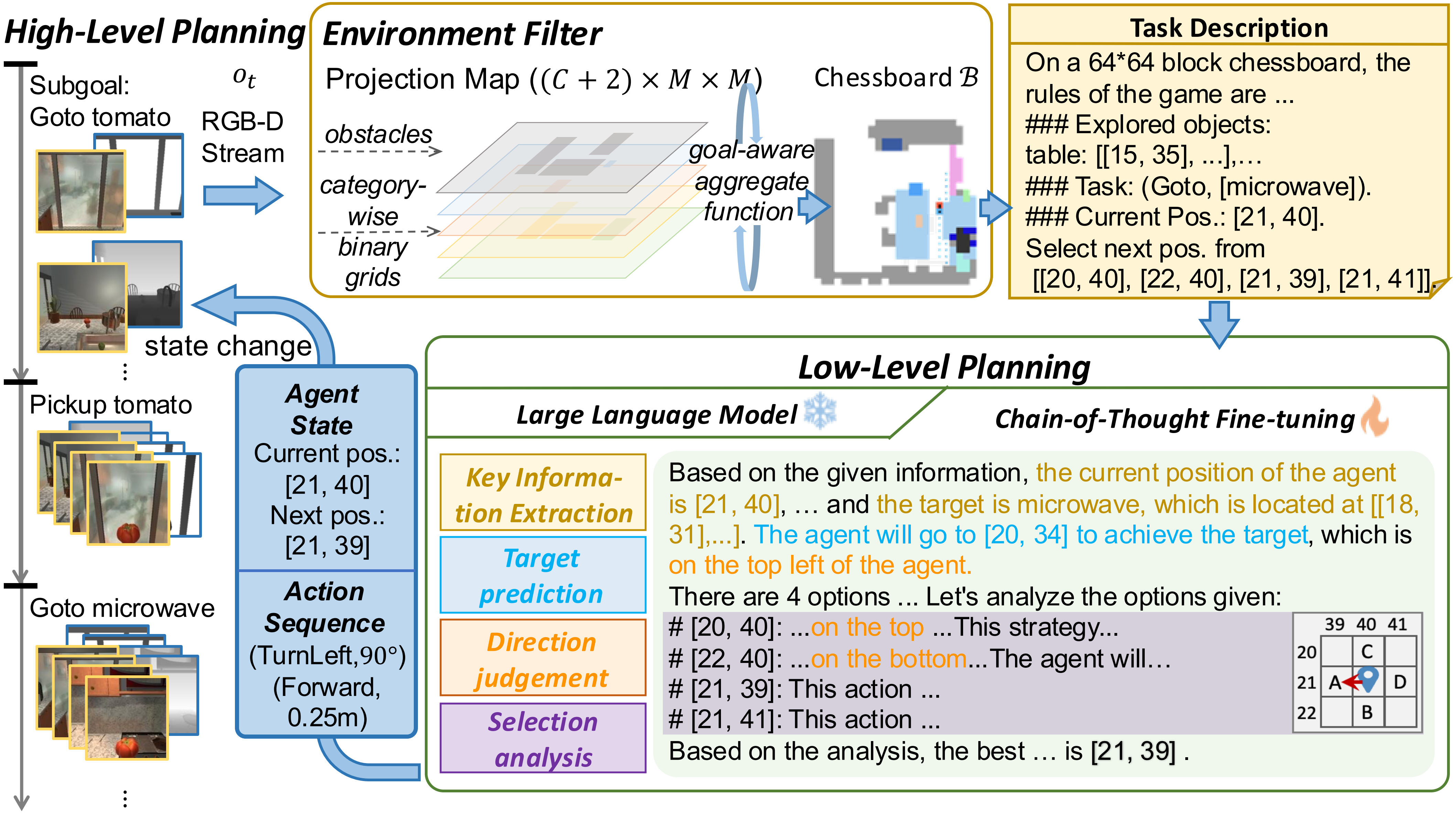
FRAMEWORK



➤ Environment Filter:

- Semantic mapping: update observation o_t to binary map $\mathcal{M}$
- Chessboard building: abstract the map into chessboard $\mathcal{B}$

FRAMEWORK

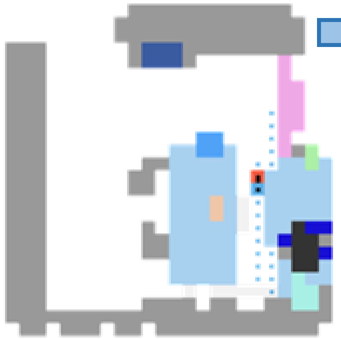


High-Level Planning
Subgoal:
Goto tomato
o_t
RGB-D
Stream
state change
Pickup tomato
Goto microwave
Environment Filter
Projection Map $((C+2) \times M \times M)$
Chessboard $\mathcal{B}$
obstacles
category-wise binary grids
goal-aware aggregate function
Task Description
On a 64*64 block chessboard, the rules of the game are ...
Explored objects:
table: [[15, 35], ...],…
Task: (Goto, [microwave]).
Current Pos.: [21, 40].
Select next pos. from
[[20, 40], [22, 40], [21, 39], [21, 41]].
Agent State
Current pos.:
[21, 40]
Next pos.:
[21, 39]
Action Sequence
(TurnLeft,90°)
(Forward, 0.25m)
Low-Level Planning
Large Language Model
Chain-of-Thought Fine-tuning
Key Information Extraction
Target prediction
Direction judgement
Selection analysis
Based on the given information, the current position of the agent is [21, 40], … and the target is microwave, which is located at [[18, 31],...]. The agent will go to [20, 34] to achieve the target, which is on the top left of the agent.
There are 4 options ... Let's analyze the options given:
[20, 40]: ...on the top ...This strategy...
[22, 40]: ...on the bottom...The agent will…
[21, 39]: This action ...
[21, 41]: This action ...
39 40 41
20 C
21 A D
22 B
Based on the analysis, the best … is [21, 39] .

CHAIN-OF-THOUGHT FINETUNING PARADIGM

- Joint Training of HLP and LLP
  - Convert expert trajectories to CoT data of 70K HLP & 195K LLP
- Chain of Thought Decision (CoT-D)

Help me heat an egg.



1.Goto fridge
2.Pickup egg
3.Goto microwave
4.Heat egg
5.Goto table
6.Put egg on table

Task Description

On a 64*64 block chessboard, the rules of the game are ...
Explored objects:
table: [[15, 35], ...],…
Task: (Goto, [microwave]).
Current Pos.: [21, 40].
Select next pos. from
[[20, 40], [22, 40], [21, 39], [21, 41]].

Based on the given information, the current position of the agent is [21, 40], … and the target is microwave, which is located at [[18, 31],...]. The agent will go to [20, 34] to achieve the target, which is on the top left of the agent.
There are 4 options ... Let's analyze the options given:
[20, 40]: ...on the top ...This strategy...
[22, 40]: ...on the bottom...The agent will…
[21, 39]: This action ...
[21, 41]: This action ...
Based on the analysis, the best … is [21, 39] .

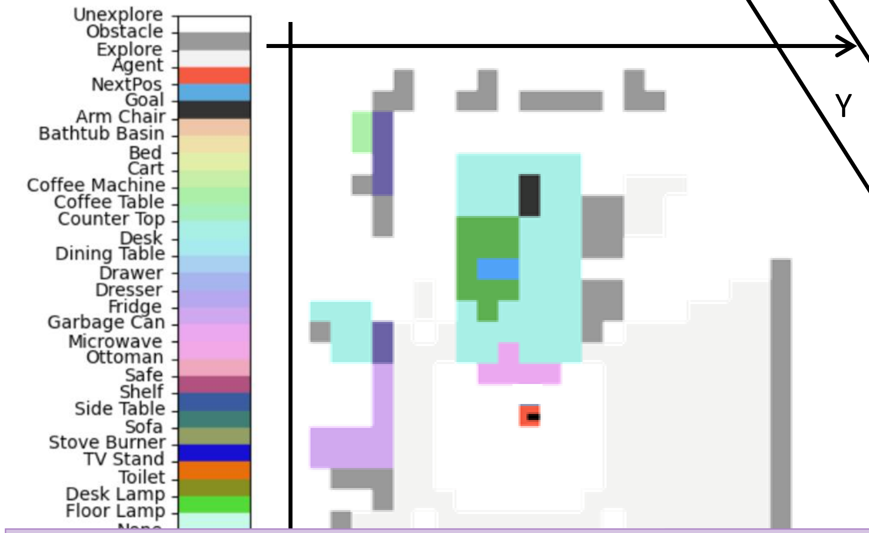
70K High-Level Planning

195K Low-Level Planning

Chain-of-Thought Decision (CoT-D)

CVPR Nashville JUNE 11-15, 2025

- Prompt System
  - Game rules R
  - Chessboard state U
  - Subgoal G
  - History actions Q



Very long context for LLMs

On a 64*64 block chessboard, the rules of the game are as follows:
Establish a coordinate system with the top left grid as (1,1). Each block can be represented by coordinates. For instance, the block in the 5th row and 3rd column is denoted as (5,3).

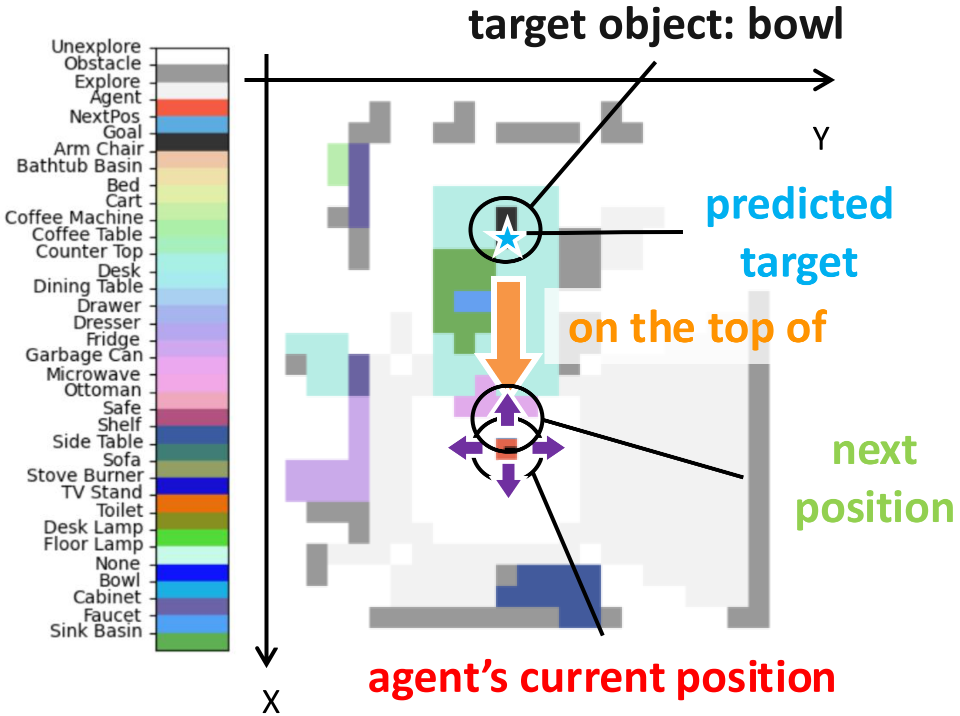
In the chessboard, there are the following explored objects:
coffeemachine: [[18, 25], [19, 25]]
cabinet: [[18, 26], [19, 26], [20, 26],...]
countertop: [[20, 30], [20, 31],...]
bowl: [[21, 33], [22, 33]]

The movement is forbidden on the object block. You can only move 1 block at a time.
Task: (GotoLocation, [bowl])
Current Position: [32, 33]
History Trajectory: []
Please select your next position from [[31, 33], [33, 33], [32, 32], [32, 34]].
Please analyze the options and choose the best one to finish the task.

Chain-of-Thought Decision (CoT-D)

CVPR Nashville JUNE 11-15, 2025

- CoT-D Subtasks
  - Key Information Extraction
  - Target Prediction
  - Direction Judgment
  - Selection Analysis



Based on the given information, **the current position of the agent is [32, 33], and the target is bowl, which is located at [[21, 33], [22, 33]]. The agent will go to [22, 33] to achieve the target, which is on the top of the agent.**

There are 4 options for the agent to move towards the target position. Now, let's analyze the options given:

[31, 33]: This position is **on the top of** the agent. By making this move, the agent approaches the target position more closely.

[33, 33]: This position is **on the bottom of** the agent. This move takes the agent away from the target position.

[32, 32]: This position is **on the left side of** the agent. This move takes the agent away from the target position.

[32, 34]: This position is **on the right side of** the agent. This move takes the agent away from the target position.

Based on the analysis above, the best option for the agent to move towards the target position is [31, 33]. Therefore, the answer is:

Next position: **[31, 33]**

EVALUATION

- Experiment Settings
  - Benchmark: ALFRED with 7 types of tasks involving both navigation and object interaction
  - LLMs: GPT-4 / LLaMA-7B / Mistral-7B

EVALUATION

- Main Results on ALFRED Benchmark

Method	Training Mode	Val Seen		Val Unseen		ΔSuccess Rate (SR)	Δ Goal-cond. SR (GC)
		SR	GC	SR	GC		
Specialists, only for ALFRED tasks							
M-TRACK	From scratch	26.70	33.21	17.29	28.98	-9.41	-4.23
FILM	From scratch	24.63	37.20	20.10	32.45	-4.53	-4.75
LEBP	From scratch	27.63	35.76	22.36	29.58	-5.27	-6.18
Generalists, based on LLMs							
SayCan	Few-shot	12.30	24.52	9.88	22.54	-2.42	-1.98
LLM-P(GPT)	Few-shot	16.45	30.11	15.36	29.88	-1.09	-0.23
R2C-GPT-4	Zero-shot	20.00	28.46	**24.00**	28.24	**+4.00**	-0.22
R2C-LLaMA-7B	Fine-tune	20.83	29.60	18.99	29.69	-1.84	**+0.09**
R2C-Mistral-7B	Fine-tune	**22.31**	**32.40**	22.35	**31.97**	+0.04	-0.43

EVALUATION

- Visualization of a trajectory in ALFRED
  - R2C-Mistral-7B

Task: **Throw both pieces of soap into the trash can.**

EVALUATION

- Ablation Study
  - CoT-D effectively improves the success rate of the model

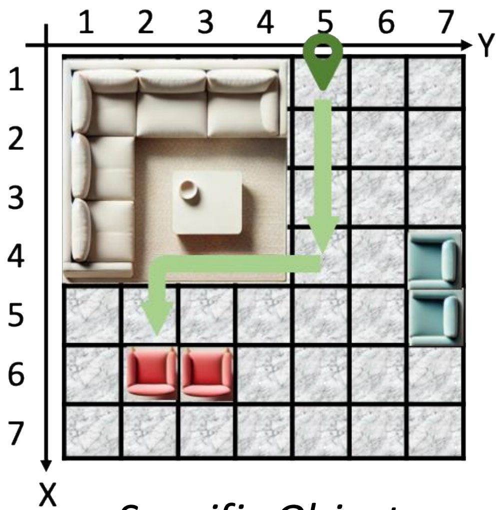
Method	Val Seen		Val Unseen	
	SR	GC	SR	GC
Base Model (R2C-Mistral-7B)	22.31	32.40	22.35	31.97
+GT Seg.	37.92	45.83	35.24	43.88
+GT Seg., GT subgoal	**48.18**	**55.13**	**53.33**	**58.18**
+GT Seg., GT subgoal – CoT-D	41.22	49.35	41.96	47.88

EVALUATION

- Open-Vocabulary Task
  - R2C can generalize to open-vocabulary task beyond object navigation

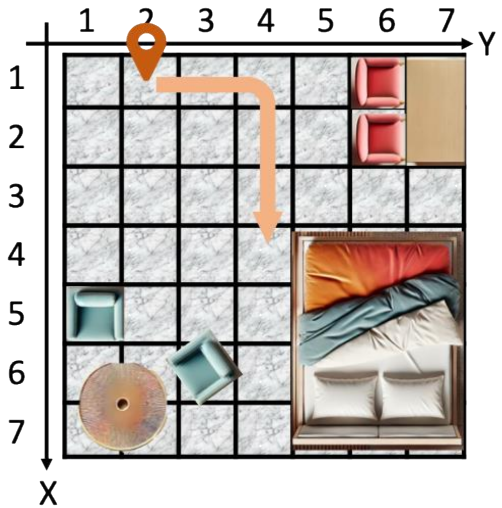
Task	Specific Obj.	Specific Loc.	Nearest Corner	Center Between	Overall
R2C-GPT-4(90°)	66.7	73.3	53.5	80.0	68.3
R2C-GPT-4(45°)	73.3	73.3	73.3	60.0	70.0

Walk to the armchair **closest to** the sofa



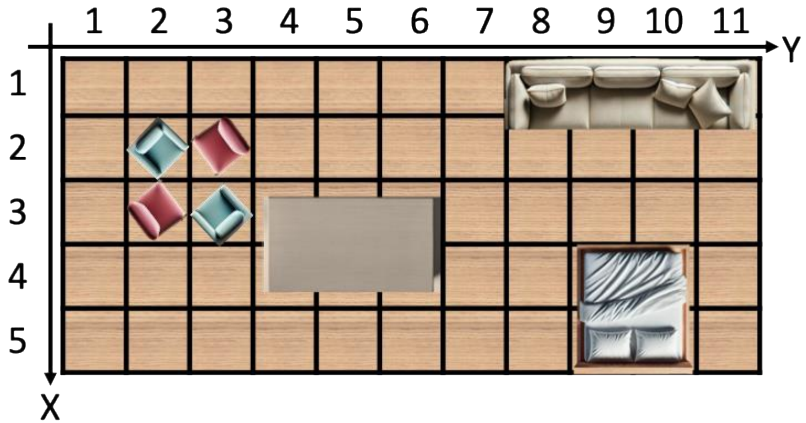
Specific Object

Go to **the center of** the room



Nearest Corner

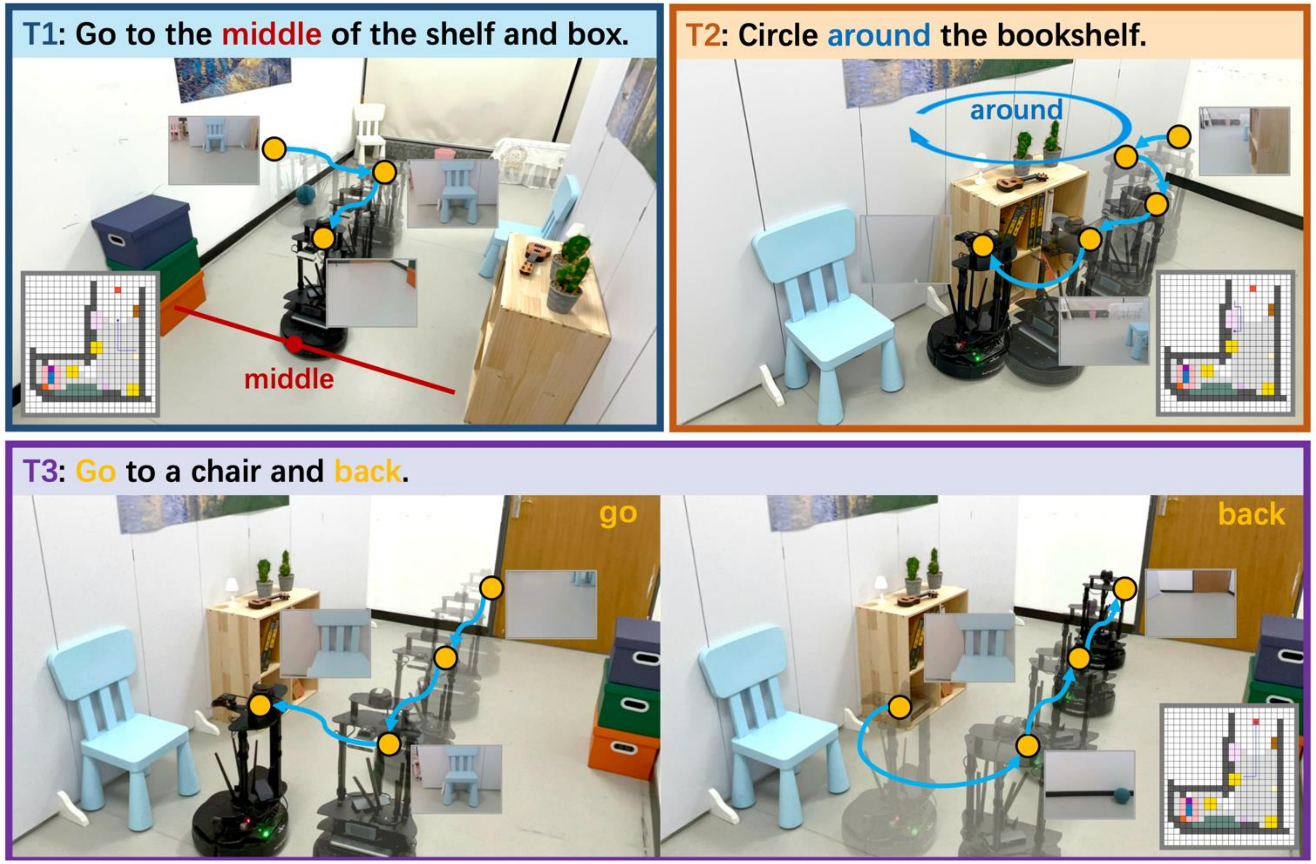
Move to the **center of the line** connecting the table and the sofa at [1,8].



Center Between

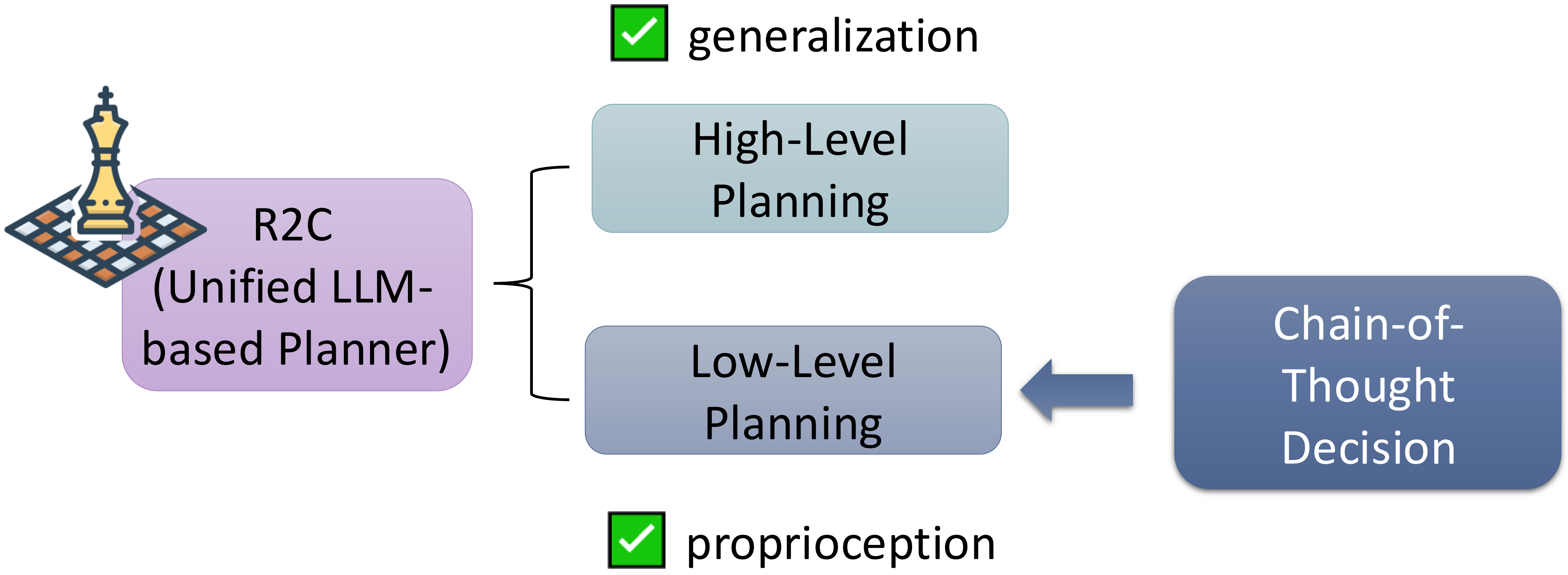
EVALUATION

- Application in Real-world Scenarios
  - Deploy R2C on LoCoBot in a 6m×6m office



CONCLUSION

- **Introduce** unified R2C framework to unlock LLM as low-level planner
- **Design** a CoT-D strategy to enhance LLM's spatial reasoning ability



Thanks for Watching and welcome to our poster!

R2C: Mapping Room to Chessboard to Unlock LLM As Low-Level Action Planner

Ziyi Bai, Hanxuan Li, Bin Fu, Chuyan Xiong, Ruiping Wang, Xilin Chen

Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China

{ziyi.bai, hanxuan.li, bin.fu, chuyan.xiong}@vipl.ict.ac.cn, {wangruiping, xlchen}@ict.ac.cn

Poster:** Sun 15 Jun 6 a.m. CST — 8 a.m. CST in **ExHall D Poster #348