JUNE 18-22, 2023
CVPR VANCOUVER, CANADA

RLLAB
http://rllab.snu.ac.kr

Meta-Explore: Exploratory Hierarchical Vision-and-Language Navigation Using Scene Object Spectrum Grounding

CVPR 2023 accepted (TUE-PM-244)

Minyoung Hwang[1], Jaeyeon Jeong[1], Minsoo Kim[3], Yoonseon Oh[2], Songhwai Oh[1]

[1]Electrical and Computer Engineering and ASRI, Seoul National University
[2]Department of Electronic Engineering, Hanyang University
[3]Interdisciplinary Major in Artificial Intelligence, Seoul National University

Contents

Vision-and-Language Navigation (VLN)

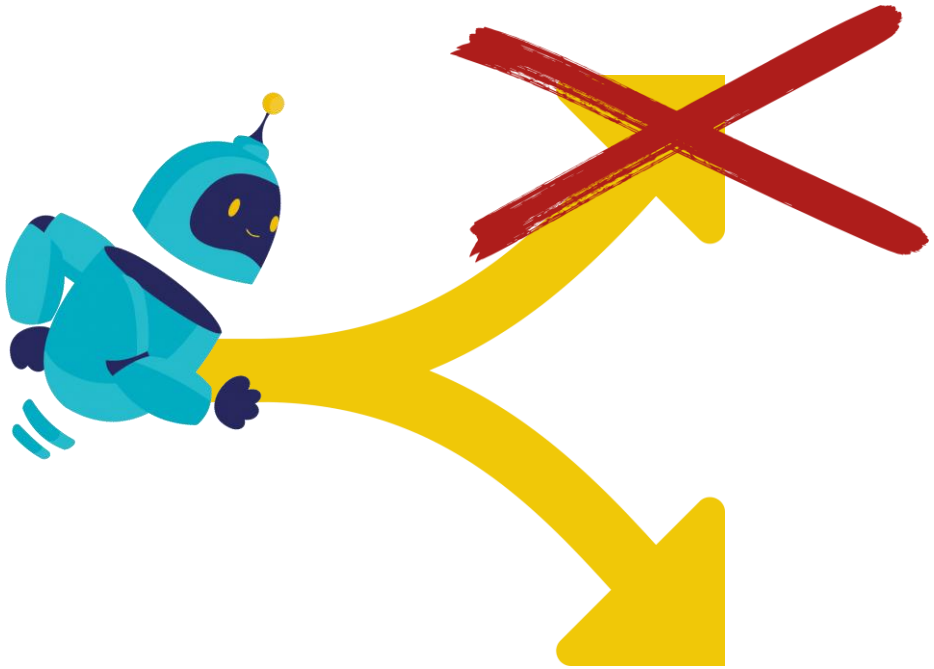
Natural Language Instruction

Visual Observation

Task: Follow the instruction and arrive at the goal location.

Vision-and-Language Navigation

If the agent is asked to **turn right** but **turns left**,
the agent may end up in **irrecoverable paths**.



*"(…) **Turn right at the end of the hall** and walk into the bedroom on your left. Stop in the doorway."*

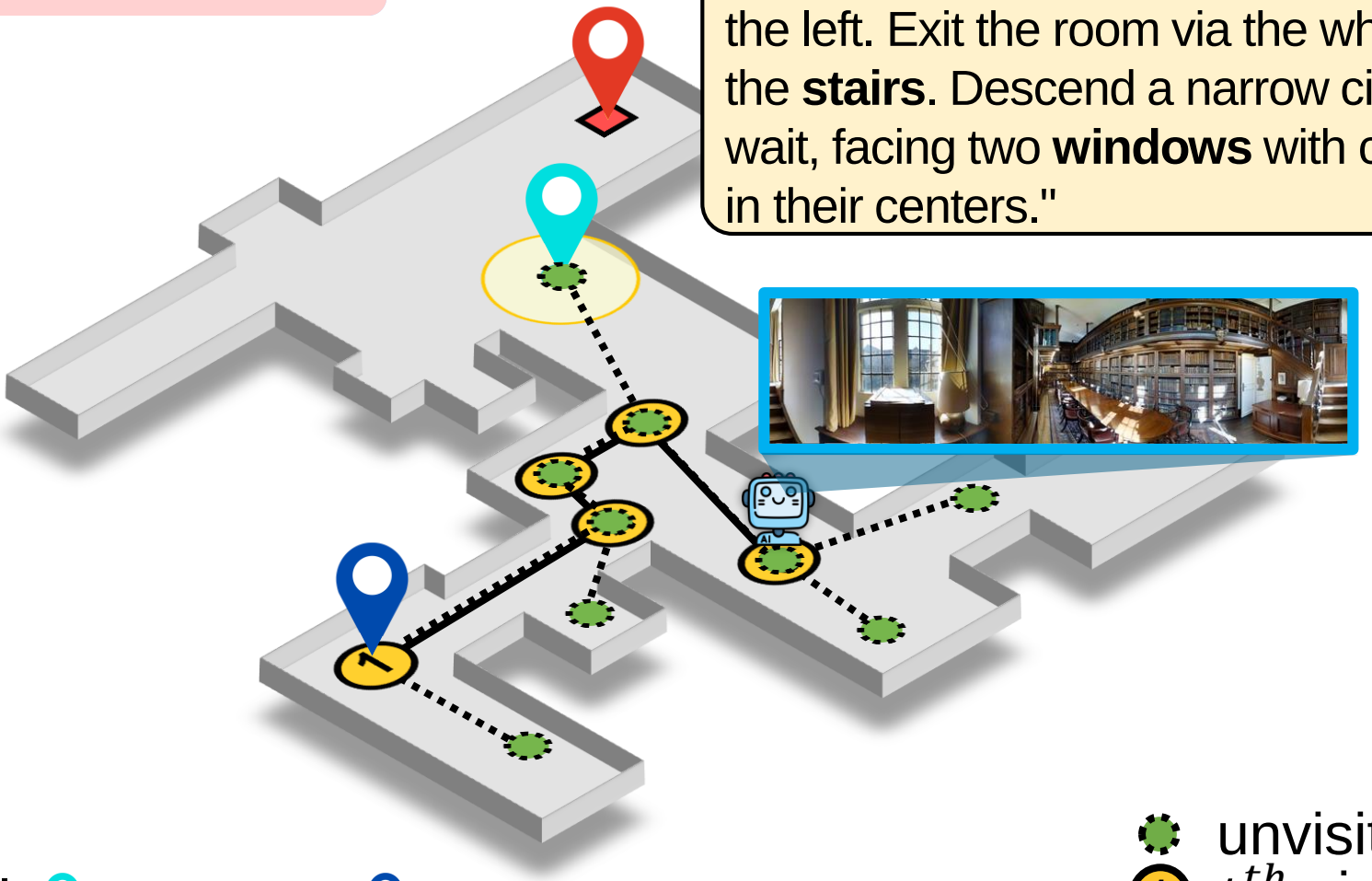
Meta-Explore

RLLAB
http://rllab.snu.ac.kr

Meta-Explore deploys an **exploitation policy** that moves the agent toward a local goal.

navigation mode: explore

Instruction: "Walk forward, keeping the long **table** to the left. Exit the room via the white **door** to the left of the **stairs**. Descend a narrow circular **stairwell** and wait, facing two **windows** with circular **stained glass** in their centers."



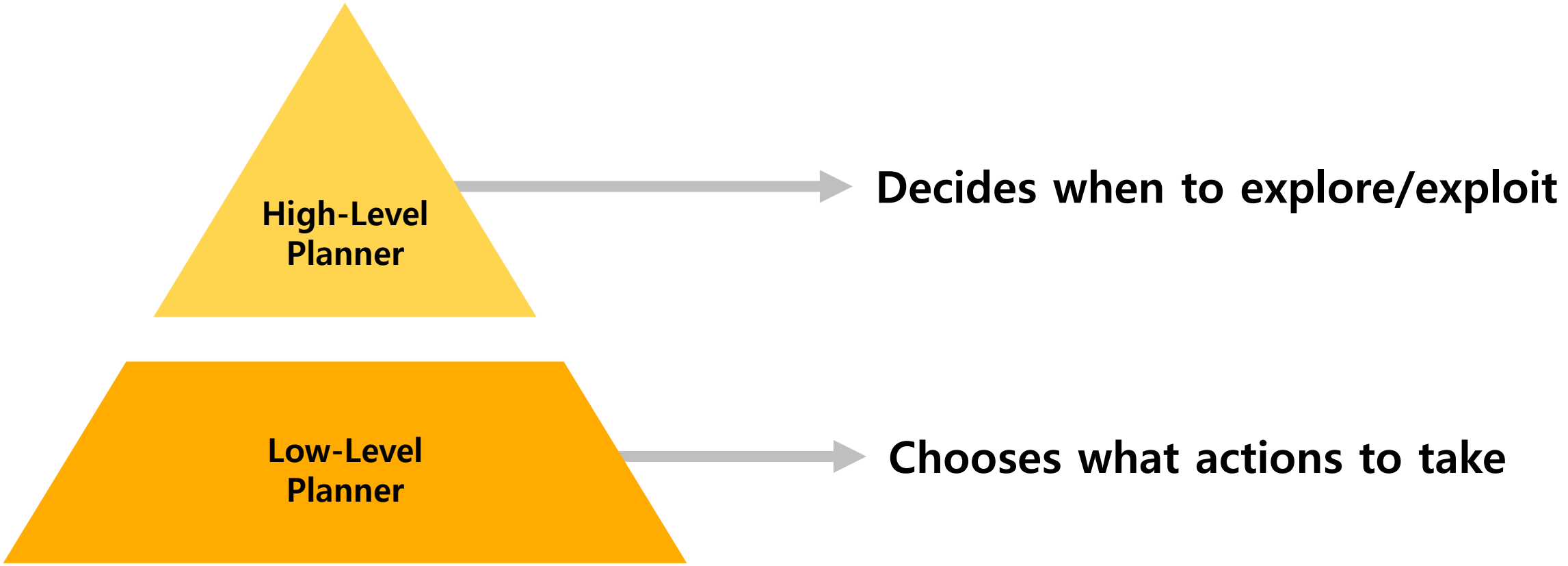
o_5

global goal local goal start

unvisited, observable node
t^{th} visited node

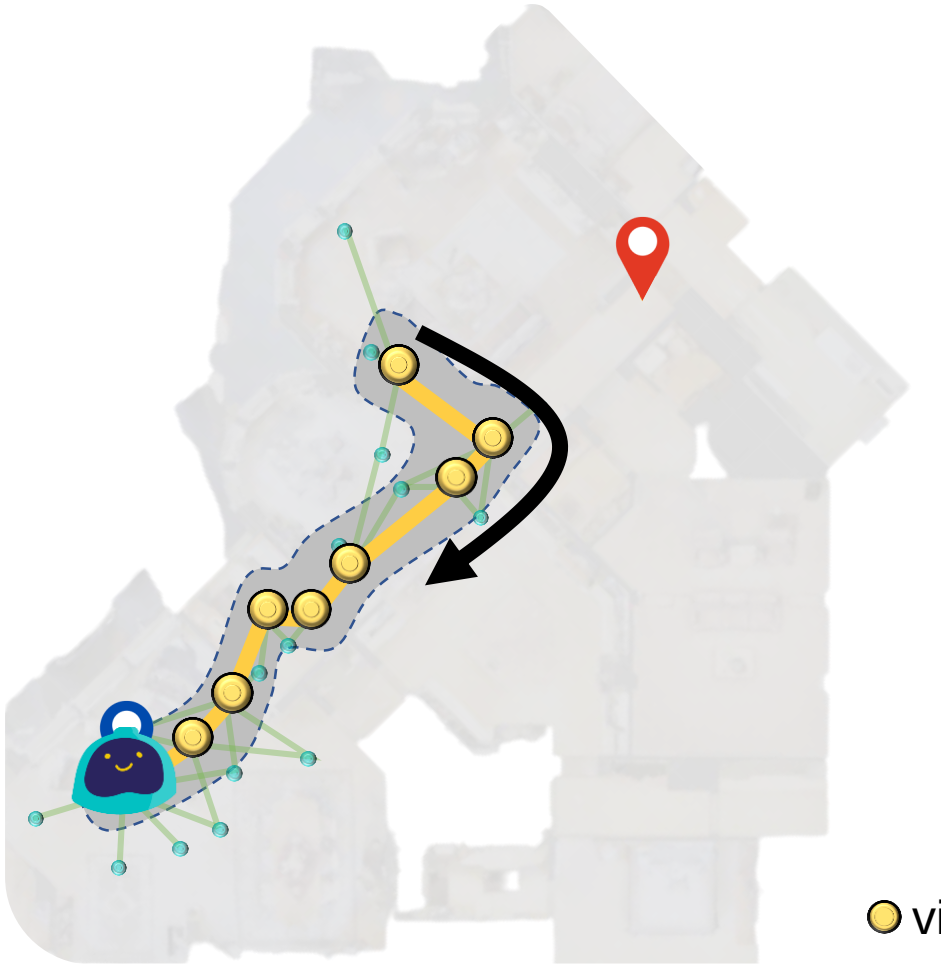
Hierarchical Exploration

Several existing studies solve this issue via hierarchical exploration.



Hierarchical Exploration

Prior work returns the agent to the last successful state.



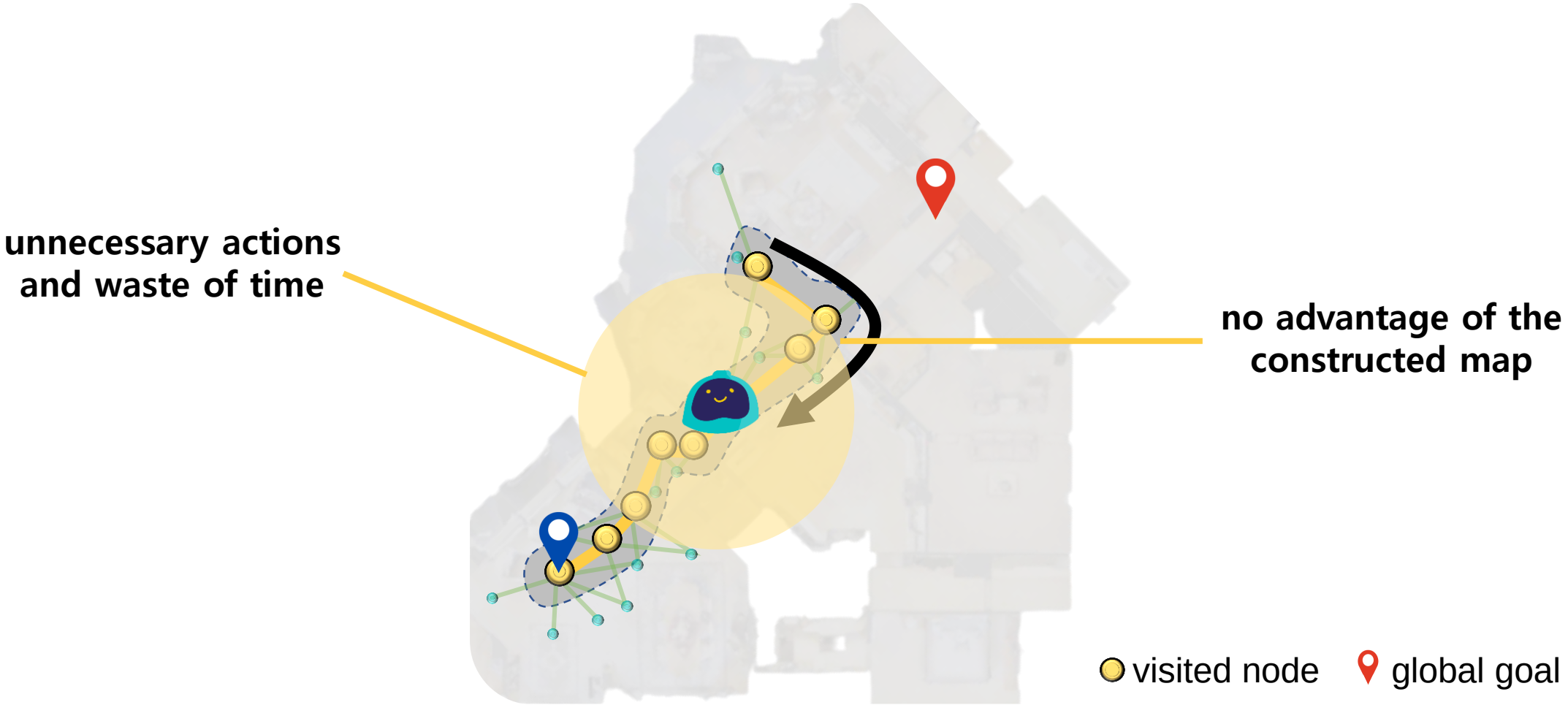
visited node global goal start

navigation mode: exploit

Hierarchical Exploration

RLLAB
http://rllab.snu.ac.kr

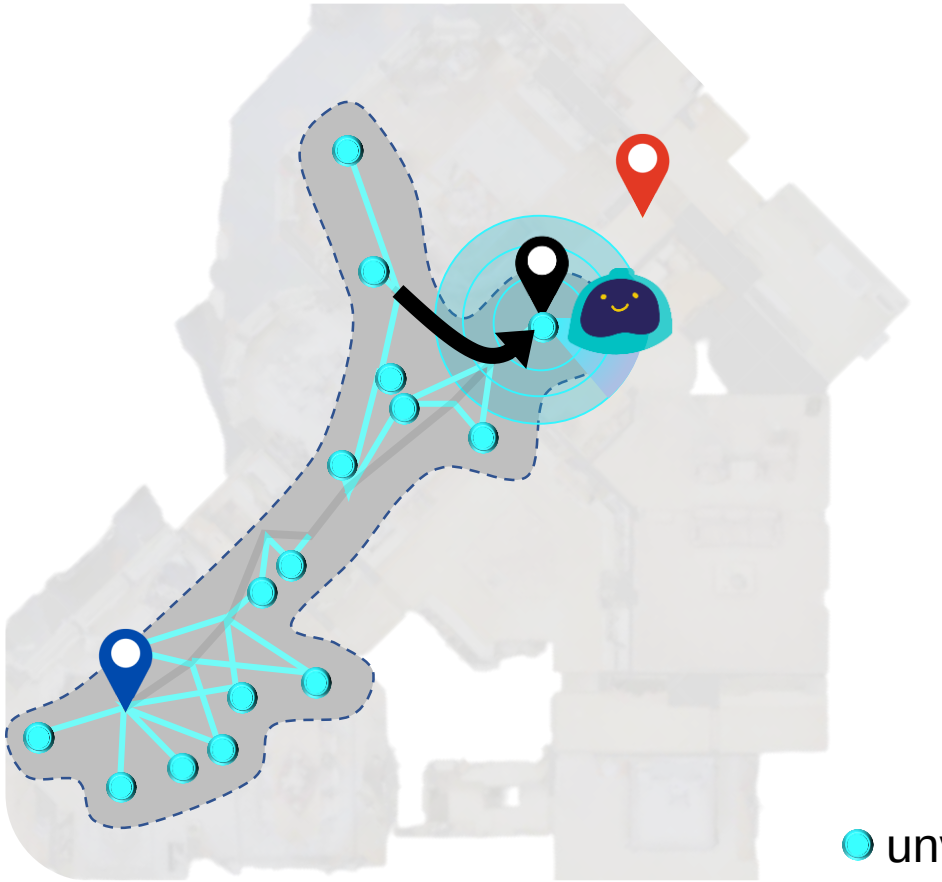
Problem of traditional exploitation methods:
The agent has to **reexplore visited regions** after the exploitation ends.



Hierarchical Exploration

Our Solution:

In the exploit mode, the agent should **move toward a well-chosen local goal** among **unvisited and observable states.**



unvisited node global goal start

Hierarchical Exploration

Our Solution:

In the exploit mode, the agent should **move toward a well-chosen local goal** among **unvisited and observable states.**

global goal local goal start
Unvisited & observable node

navigation mode: exploit

Hierarchical Exploration

RLLAB
http://rllab.snu.ac.kr

Our Solution:

In the exploit mode, the agent should **move toward a well-chosen local goal** among **unvisited and observable states.**

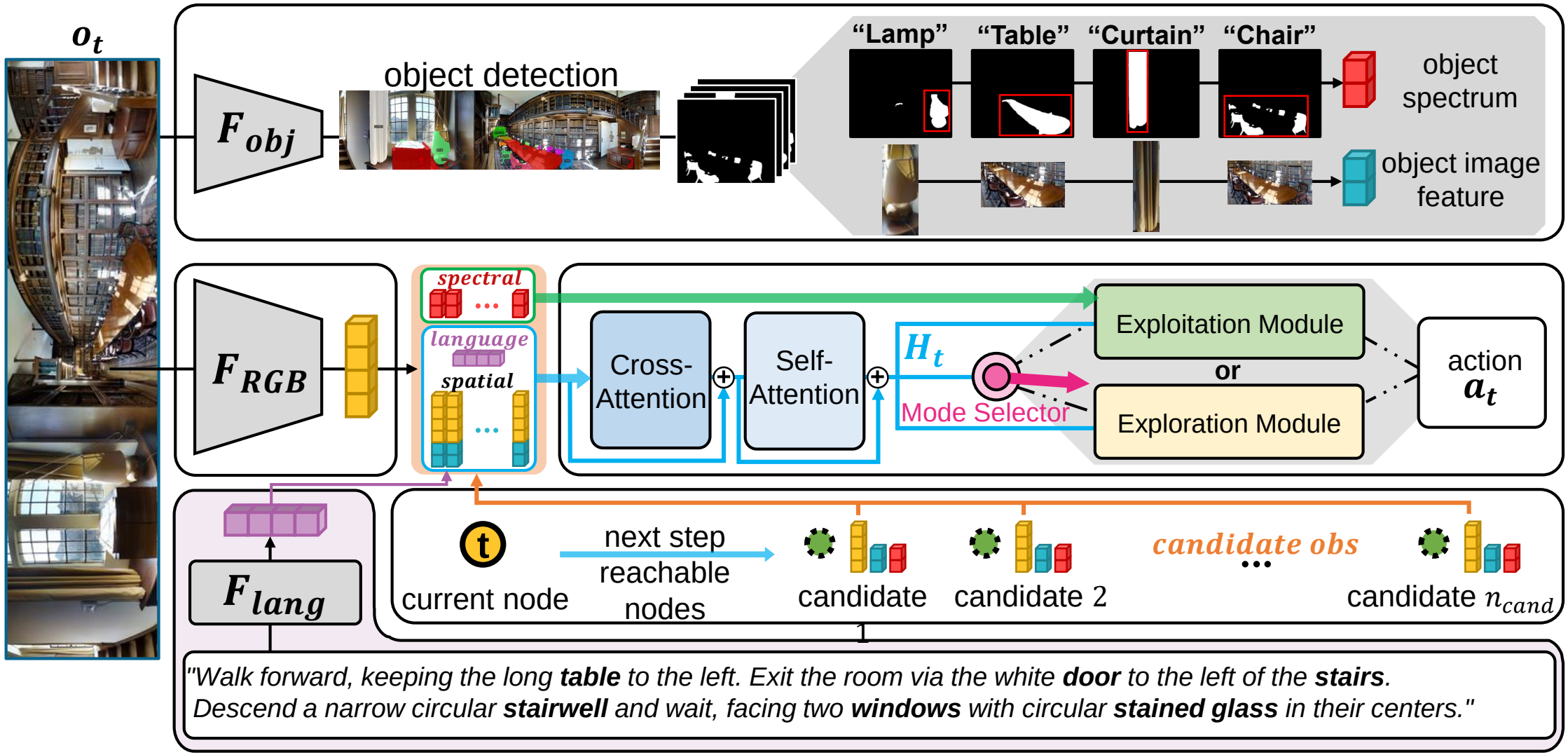
higher chance to get closer to the global goal

global goal local goal start

Unvisited & observable node

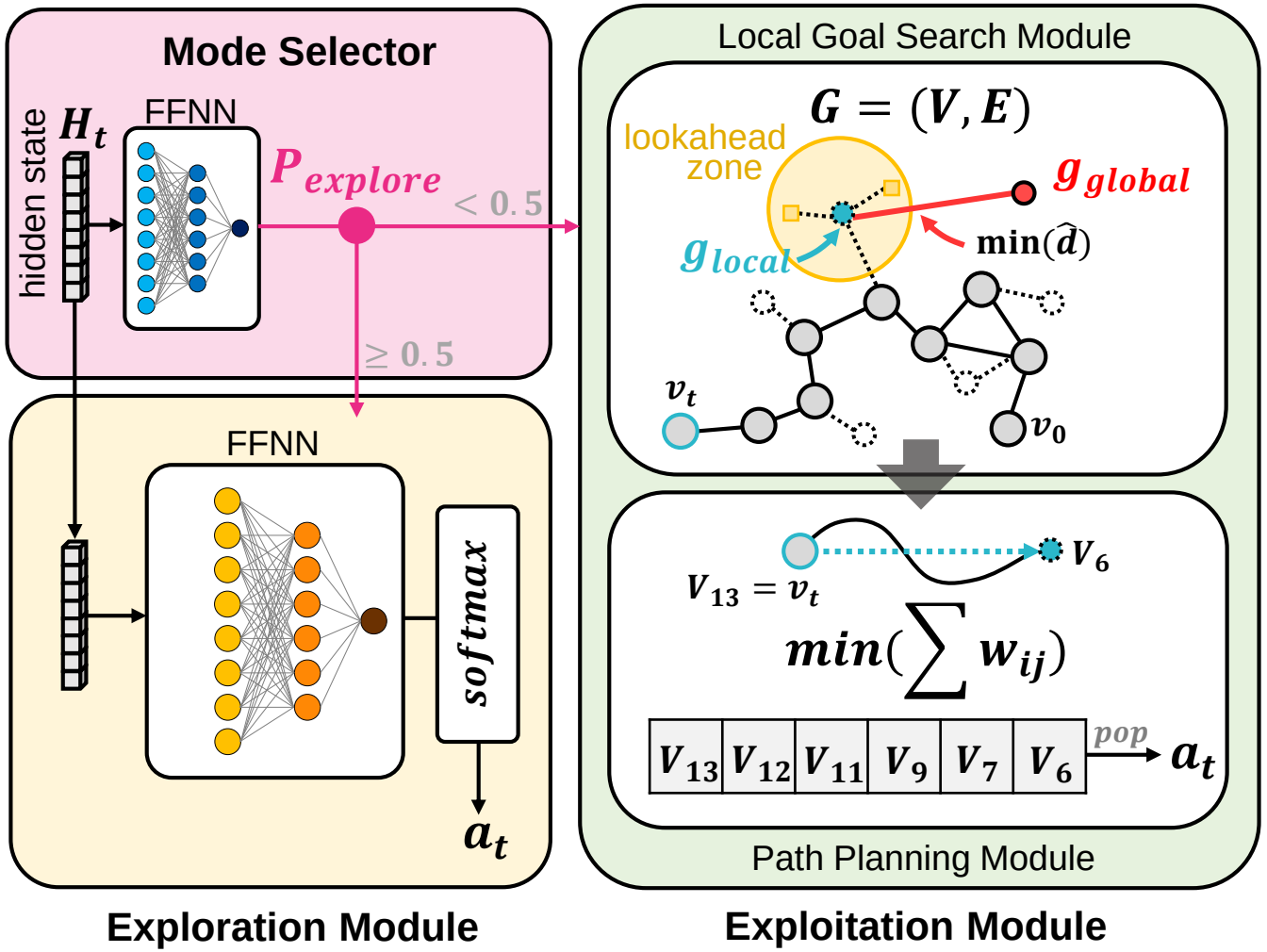
Meta-Explore

We propose a learnable hierarchical exploration method called **Meta-Explore**, which decides (1) when to explore or exploit and (2) a new imagined local goal to seek during exploitation.



Meta-Explore

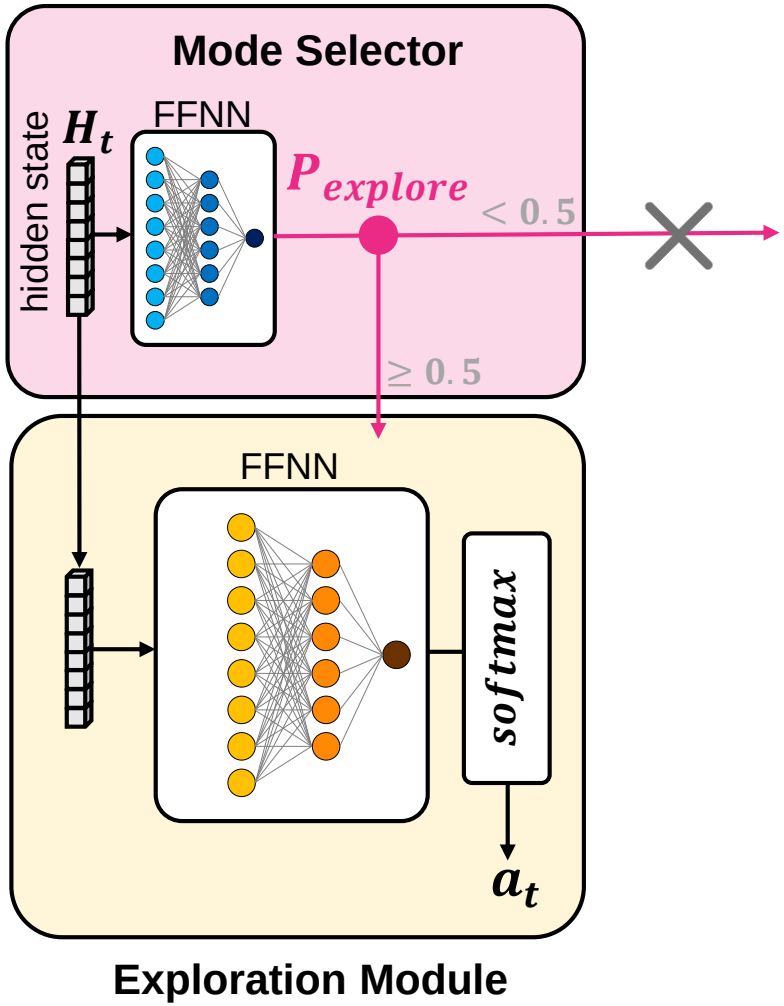
Meta-Explore consists of a **mode selector** and **two navigation modules** corresponding to two modes: **exploration** and **exploitation**.



Meta-Explore

At each timestep, the **mode selector** chooses to **explore or exploit**.
The explore-exploit switching decision occurs by **estimating the probability to explore**.

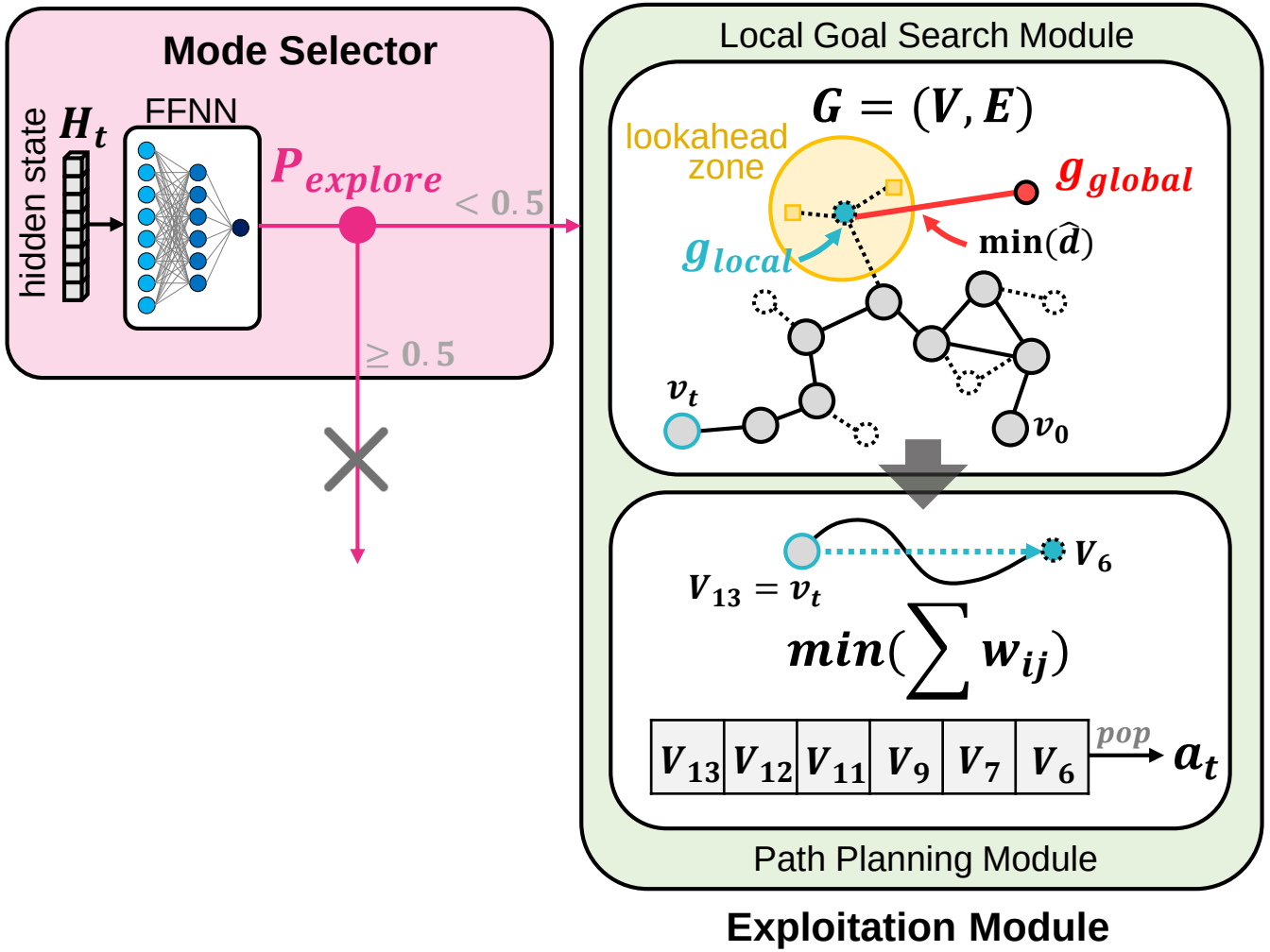
If
$P_{explore} \geq 0.5$



Meta-Explore

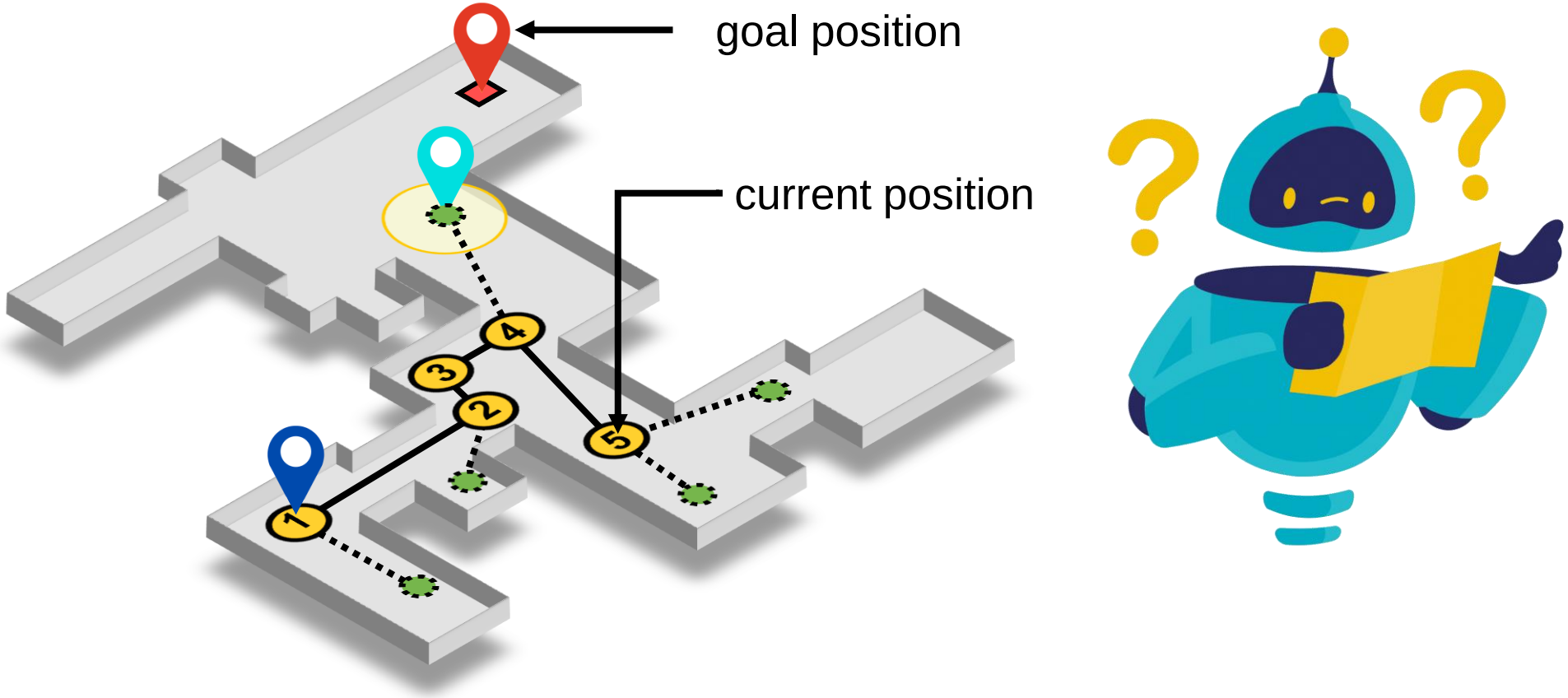
At each timestep, the **mode selector** chooses to **explore or exploit**.
The explore-exploit switching decision occurs by **estimating the probability to explore**.

If
$P_{explore} < 0.5$



Meta-Explore

When the mode selector recognizes that **the agent is not following the instruction** successfully, the **mode is switched** to exploitation.

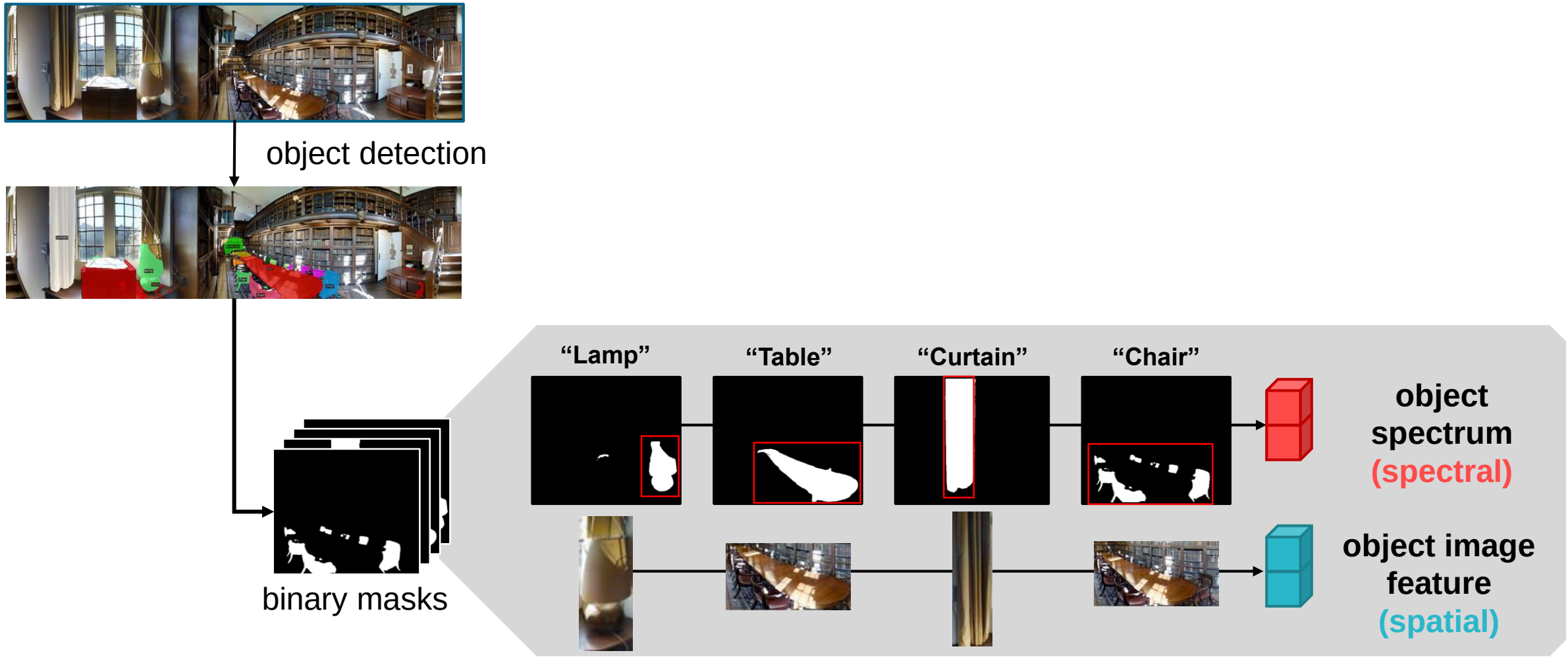


In the **exploitation mode**, the agent seeks a new local goal with the highest correspondence against the language instructions from the previously unvisited candidate nodes.

Meta-Explore

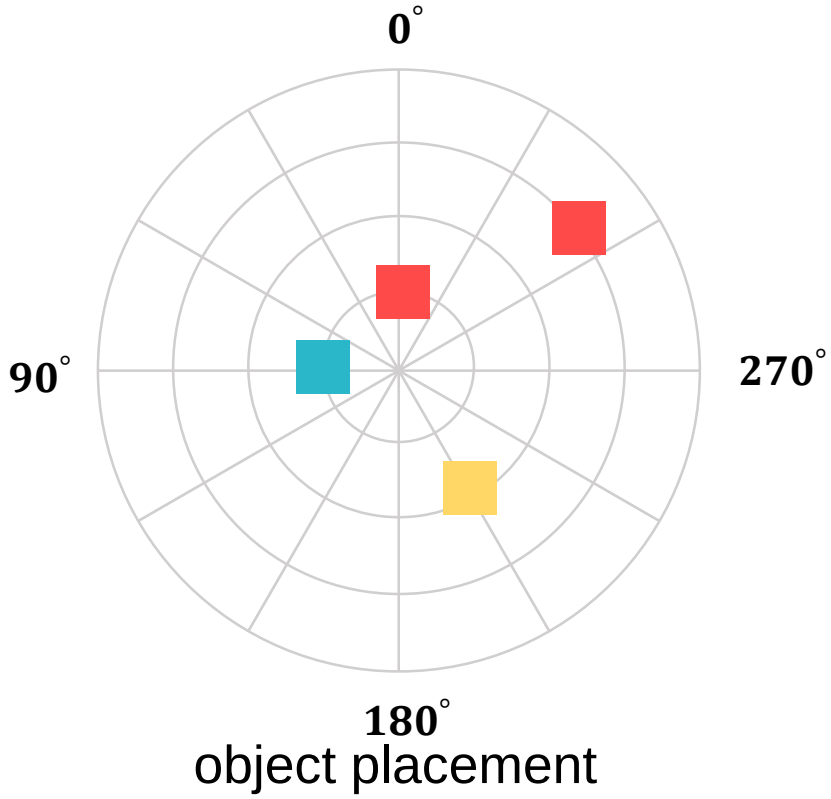
RLLAB
http://rllab.snu.ac.kr

Meta-Explore imagines regretful explorations with **semantically meaningful clues**.



Meta-Explore

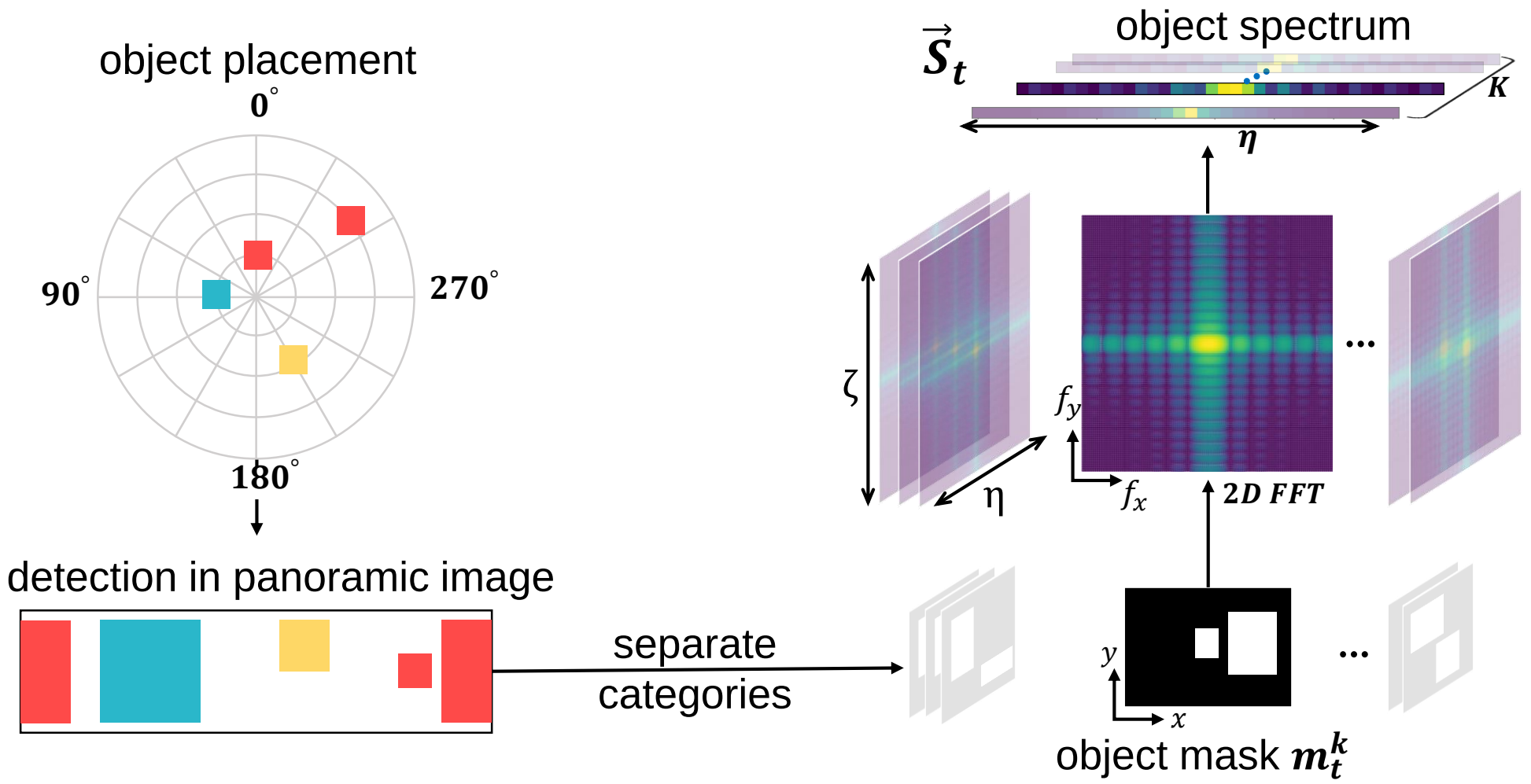
The key to our exploitation approach is understanding the **object placements** in **spectral-domain**.



"What objects are placed around the agent at what frequencies?"

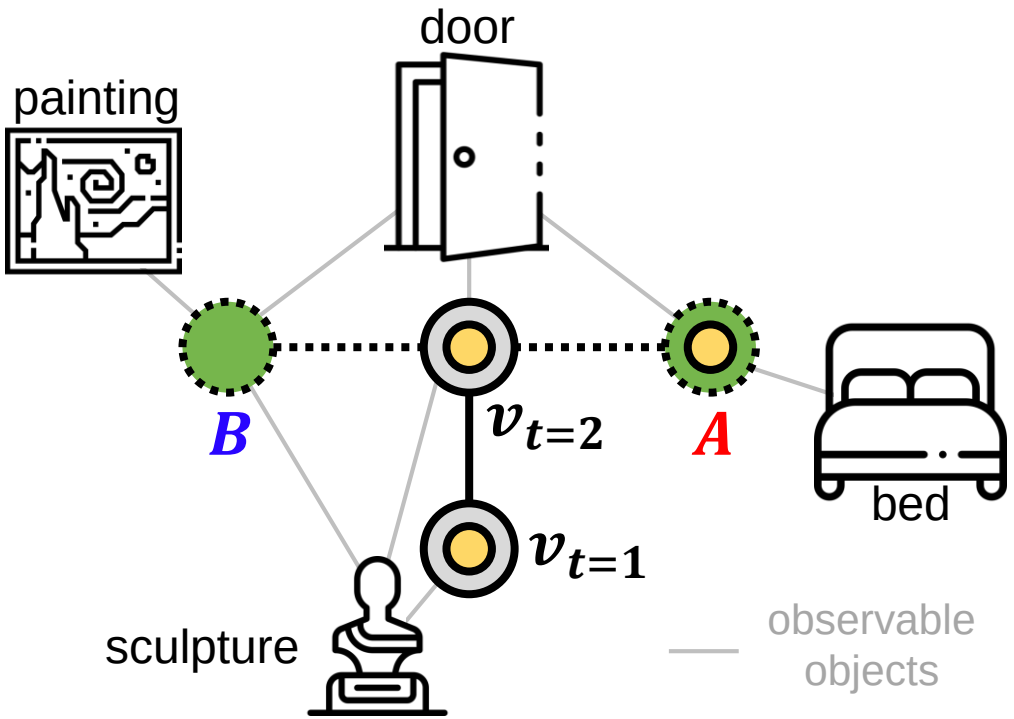
Scene Object Spectrum (SOS)

We present a **novel** visual representation, called **scene object spectrum (SOS)**, which performs **category-wise 2D Fourier transform** of **detected objects**.



Toy Example

Combining the exploitation policy and SOS features, the agent can correct its path by choosing a **promising local goal**.



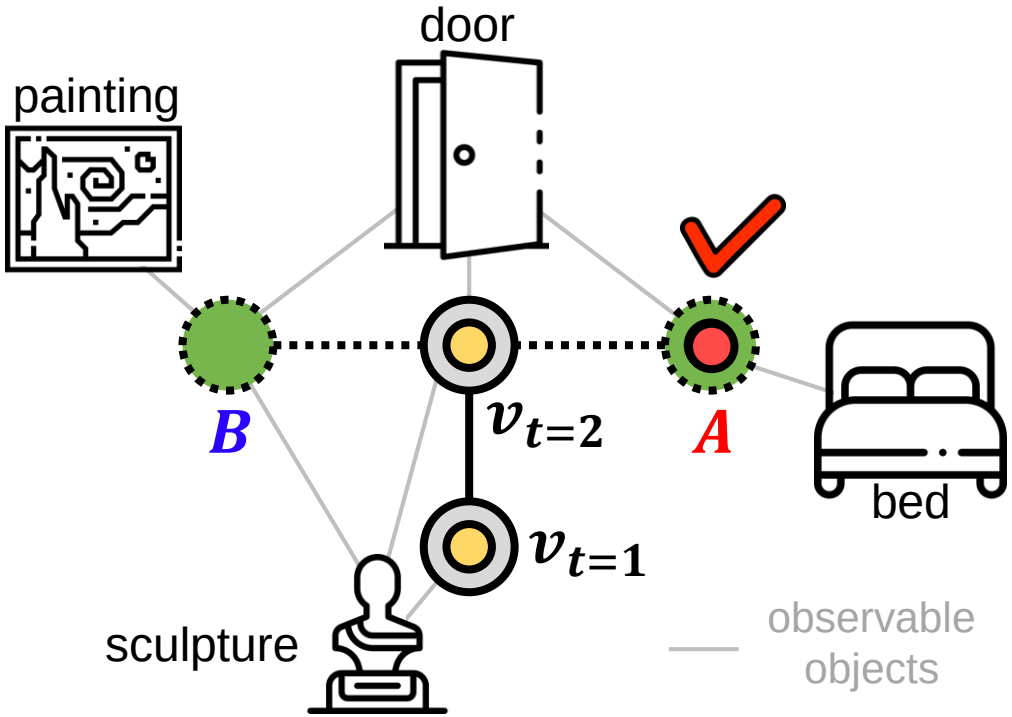
$$S_{nav}(\mathcal{T}') = \frac{\sum_{i=1}^{B}\sum_{j=1}^{t'}\left(\frac{\hat{\delta}(w_i^o)}{|\hat{\delta}(w_i^o)|}\cdot\frac{\vec{S}'_j}{|\vec{S}'_j|}\right)\left((\hat{\delta}(w_i^o)-\hat{\delta}(\overline{w}^o))\cdot(\vec{S}'_j-\overline{\vec{S}}')\right)}{\sqrt{\frac{t'}{B}\cdot\sum_{i=1}^{B}(\hat{\delta}(w_i^o)-\hat{\delta}(\overline{w}^o))^2\sum_{j=1}^{t'}(\vec{S}'_j-\overline{\vec{S}}')^2}}$$

navigation score S_{nav}

We score the corrected trajectories to measure the alignment with the language instruction L.

Toy Example

Suppose the agent is in the **exploitation mode**.
The agent wants to choose the **local goal** between A and B.



trajectory-instruction similarity matrices

	w_1^o	w_2^o	w_3^o
$v_{t=1}$	0.9	0.0	0.0
$v_{t=2}$	0.3	1.0	0.0
$v_{t=3}$	0.0	0.0	1.0

$v_{t=3} = A,\ S_{nav} = 0.296$

	w_1^o	w_2^o	w_3^o
$v_{t=1}$	0.9	0.0	0.0
$v_{t=2}$	0.3	1.0	0.0
$v_{t=3}$	0.5	0.4	0.0

$v_{t=3} = B,\ S_{nav} = 0.201$

Experiment Results: R2R

RLLAB
http://rllab.snu.ac.kr

Meta-Explore improves success rate and SPL compared to hierarchical baselines.

Methods	Memory	Exploit	Val Seen				Val Unseen				Test Unseen			
			SR↑	SPL↑	TL↓	NE↓	SR↑	SPL↑	TL↓	NE↓	SR↑	SPL↑	TL↓	NE↓
Random	-	-	16	-	9.58	9.45	16	-	9.77	9.23	13	12	9.89	9.79
Human	-	-	-	-	-	-	-	-	-	-	11.85	1.61	86	76
Seq2Seq [2]	Rec	✗	6.0	39	11.33	-	22	-	**8.39**	7.84	20	18	**8.13**	7.85
VLN⟳BERT [32]	Rec	✗	72	68	**11.13**	2.90	63	57	12.01	3.93	63	57	12.35	4.09
SMNA† [21]	Rec	**homing**	69	63	11.69	3.31	47	41	12.61	5.48	61	56	-	4.48
Regretful-Agent [22]	Rec	**homing**	69	63	-	3.23	50	41	-	5.32	48	40	-	5.69
FAST (short) [35]	Rec	**homing**	-	-	-	-	56	43	21.17	4.97	54	41	22.08	5.14
FAST (long) [35]	Rec	**homing**	70	04	188.06	3.13	63	02	224.42	4.03	61	03	196.53	4.29
HAMT-e2e [34]	Seq	✗	76	72	11.15	2.51	66	61	11.46	**2.29**	65	60	12.27	3.93
DUET [24]	Top. Map	✗	79	73	12.32	2.28	**72**	60	13.94	3.31	69	59	14.73	3.65
SSM [26]	Top. Map	**jump**	71	62	14.7	3.10	62	45	20.7	4.32	61	46	20.4	4.57
Meta-Explore (Ours)	Top. Map	**local goal**	**81**	**75**	11.95	**2.11**	**72**	**62**	13.09	3.22	**71**	**61**	14.25	**3.57**

Table 1. Comparison and evaluation results of the baselines and our model in the R2R Navigation Task.
Gray shaded rows describe hierarchical navigation baselines. Three memory types: Rec(recurrent), Seq(sequential), and Top. Map(topological map)

Experiment Results: SOON

Meta-Explore outperforms other baselines and shows significant **generalization performance**.

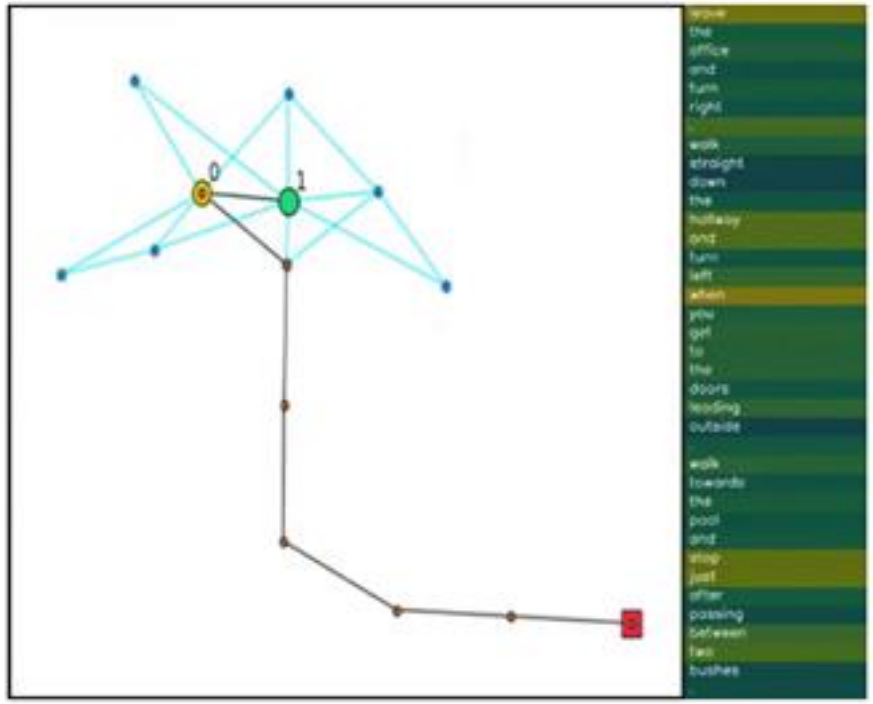
Methods	Memory	Exploit	Val Seen Instruction				Val Seen House				Test Unseen House			
			SR↑	SPL↑	OSR↑	FSPL↑	SR↑	SPL↑	OSR↑	FSPL↑	SR↑	SPL↑	OSR↑	FSPL↑
Human	-	-	-	-	-	-	-	-	-	-	90.4	59.2	91.4	51.1
Random	Rec	✗	0.0	1.5	0.1	1.4	0.1	0.0	0.4	0.9	2.1	0.4	2.7	0.0
Speaker-Follower [28]	Rec	✗	97.9	97.7	97.8	24.5	61.2	60.4	69.4	**9.1**	7.0	6.1	9.8	0.6
RCM [49]	Rec	✗	84.0	82.6	89.1	10.9	62.4	60.9	72.7	7.8	7.4	6.2	12.4	0.7
AuxRN [23]	Rec	✗	98.4	97.4	**98.7**	13.7	68.8	**67.3**	**78.5**	8.3	8.1	6.7	11.0	0.5
GBE w/o GE	Top. Map	✗	89.5	88.3	91.8	24.2	62.5	60.8	73.0	6.7	11.4	8.7	18.8	0.8
GBE [16]	Top. Map	✗	98.4	97.9	98.6	**44.2**	**76.3**	62.5	64.1	7.3	11.9	10.2	19.5	1.4
GBE†	Top. Map	✗	-	-	-	-	19.5	13.3	28.5	1.2	12.9	9.2	21.5	0.5
DUET [24]	Top. Map	✗	94.0	91.6	90.0	31.1	36.3	22.6	50.9	3.8	33.4	21.4	43.0	**4.2**
Meta-Explore (Ours)	Top. Map	**local goal**	**100.0**	**99.1**	96.0	33.9	44.7	34.8	52.7	8.9	**39.1**	**25.8**	**48.7**	4.0

Table 2. Comparison and evaluation results of the baselines and our model in the SOON Navigation Task.

Experiment Visualization 1: R2R

RLLAB
http://rllab.snu.ac.kr

Meta-Explore constructs a topological map during exploration and uses it for decision making.

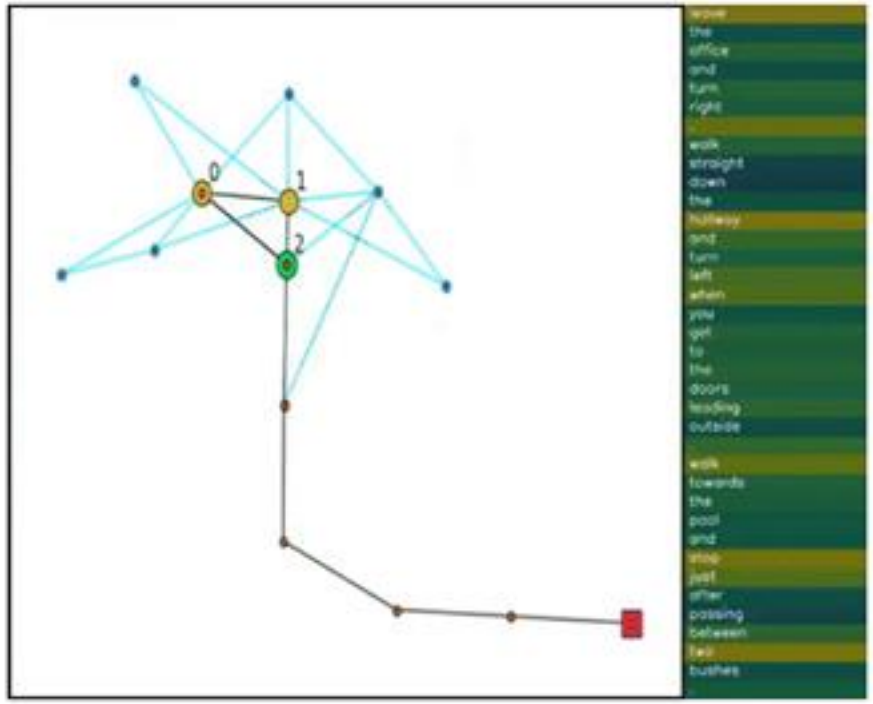


visited node visited node visited node visited node shortest path

Experiment Visualization 1: R2R

RLLAB
http://rllab.snu.ac.kr

Meta-Explore constructs a topological map during exploration and uses it for decision making.

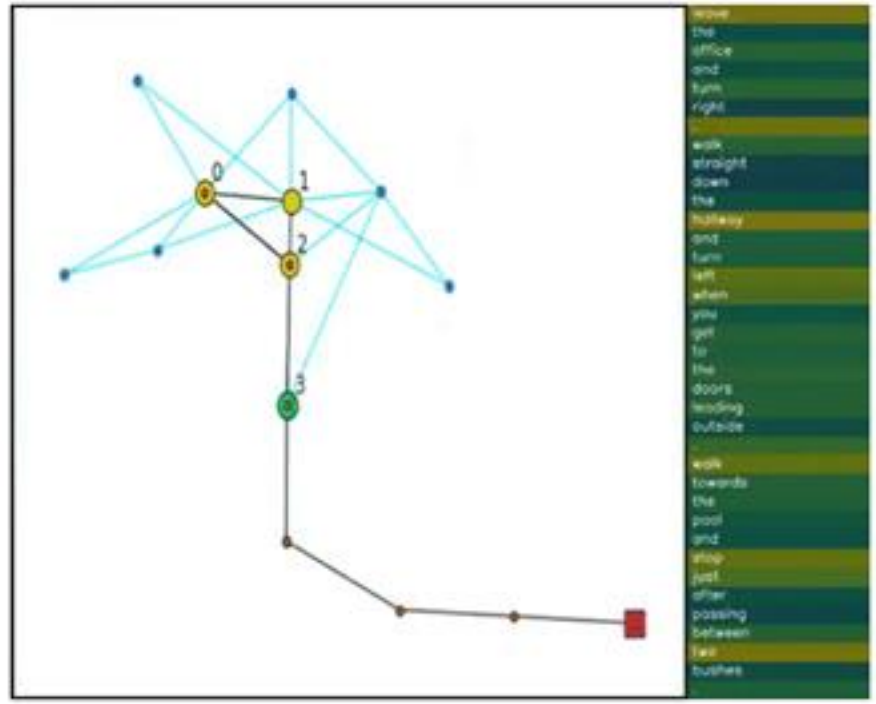


visited node visited node visited node visited node shortest path

Experiment Visualization 1: R2R

RLLAB
http://rllab.snu.ac.kr

Meta-Explore constructs a topological map during exploration and uses it for decision making.

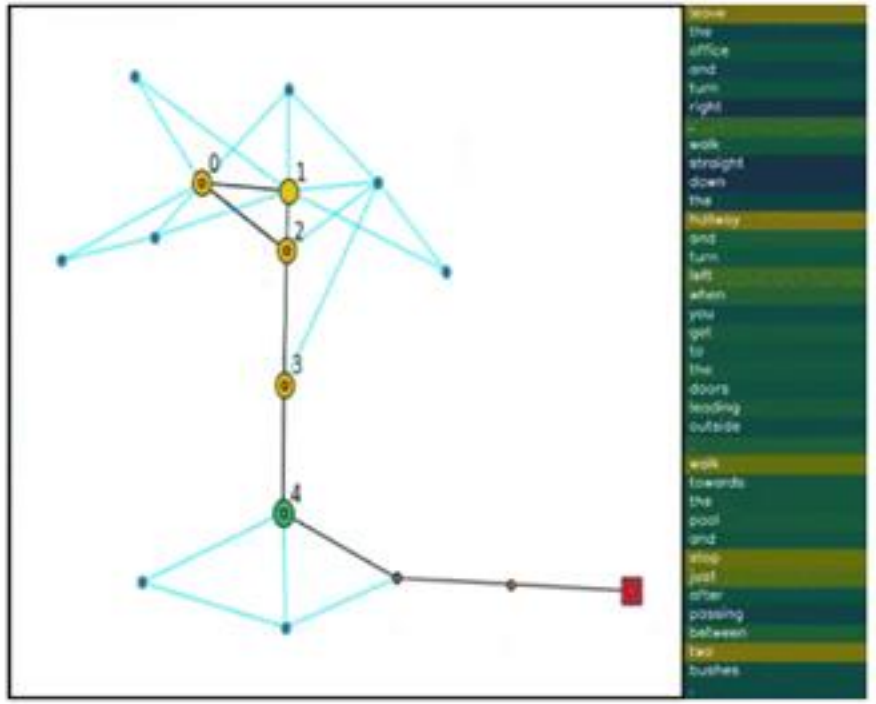


visited node visited node visited node visited node shortest path

Experiment Visualization 1: R2R

RLLAB
http://rllab.snu.ac.kr

Meta-Explore constructs a topological map during exploration and uses it for decision making.

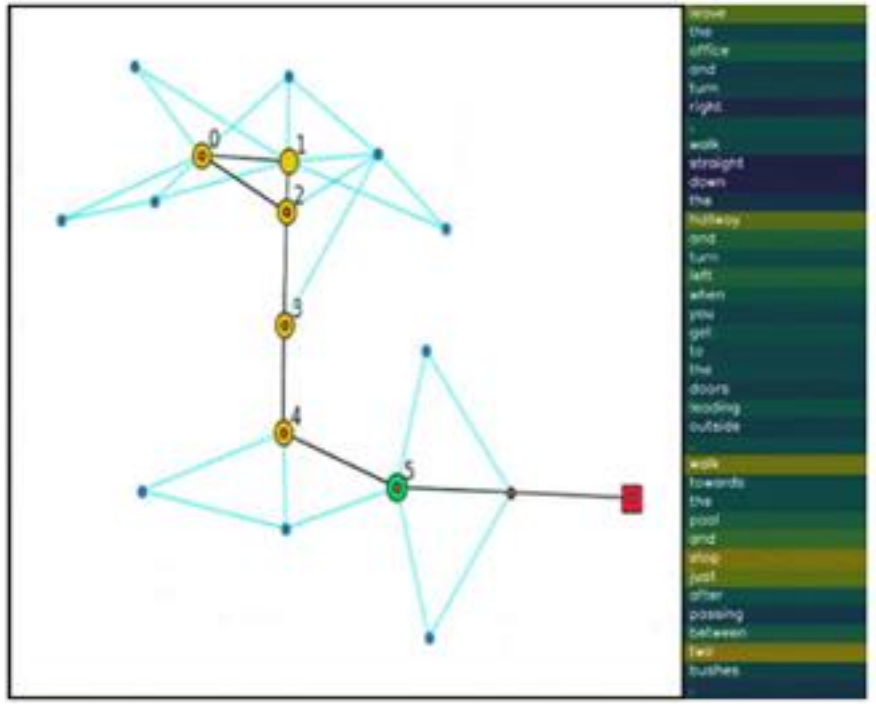


visited node visited node visited node visited node shortest path

Experiment Visualization 1: R2R

RLLAB
http://rllab.snu.ac.kr

Meta-Explore constructs a topological map during exploration and uses it for decision making.

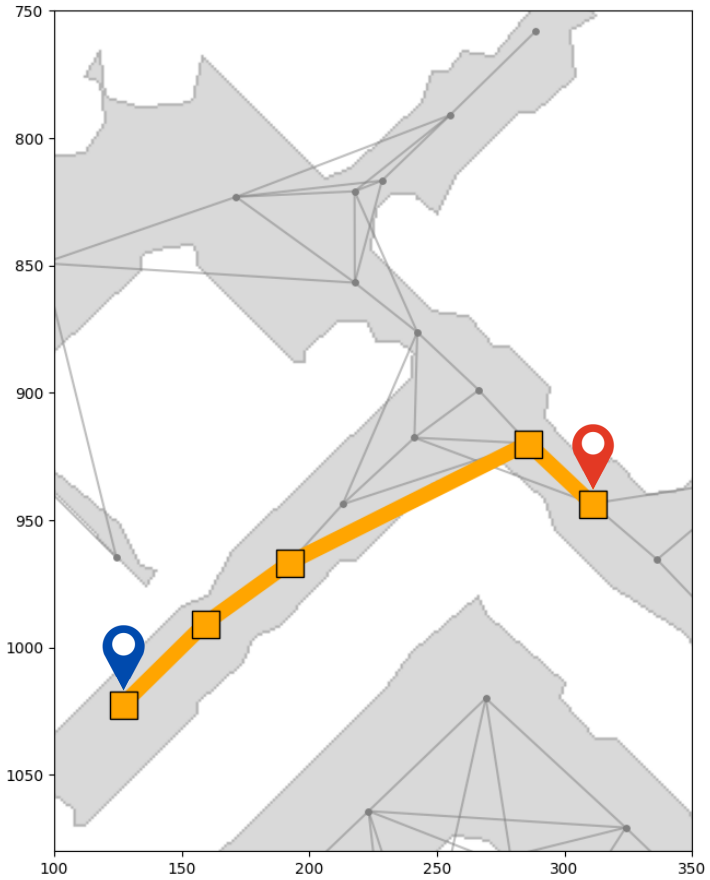


visited node visited node visited node visited node shortest path

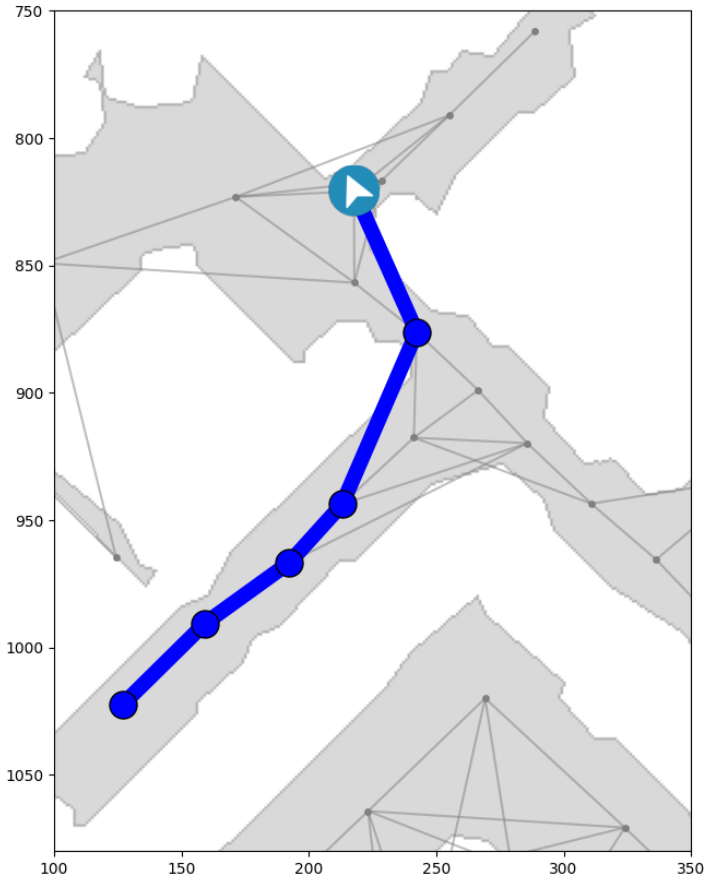
Experiment Visualization 2: Local Goal Search

RLLAB
http://rllab.snu.ac.kr

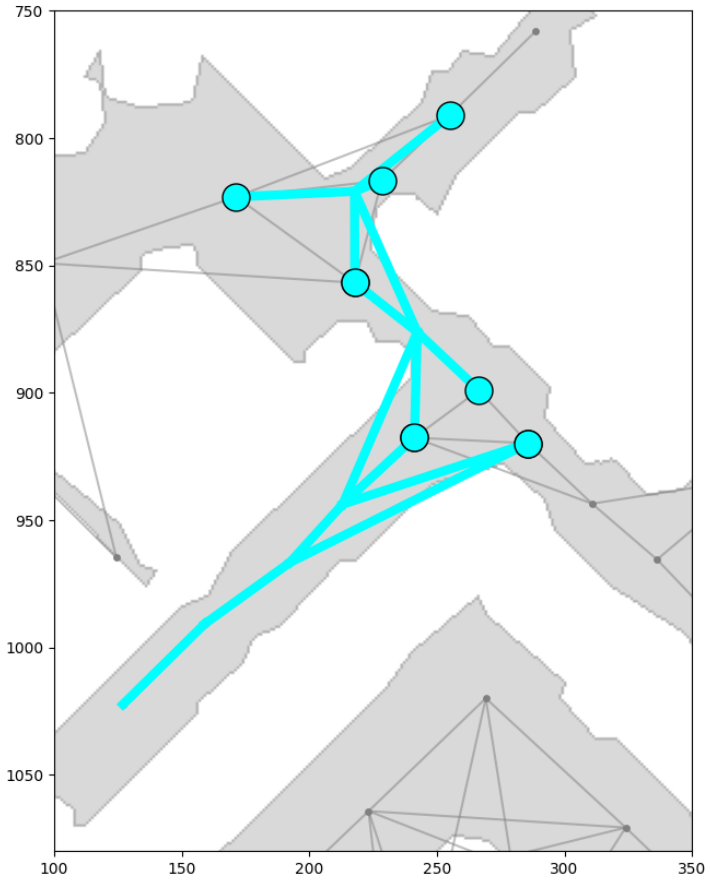
visited node
unvisited node
ground truth trajectory
global goal
local goal
start
agent
navigation graph



Ground Truth Traj.



Current Traj.



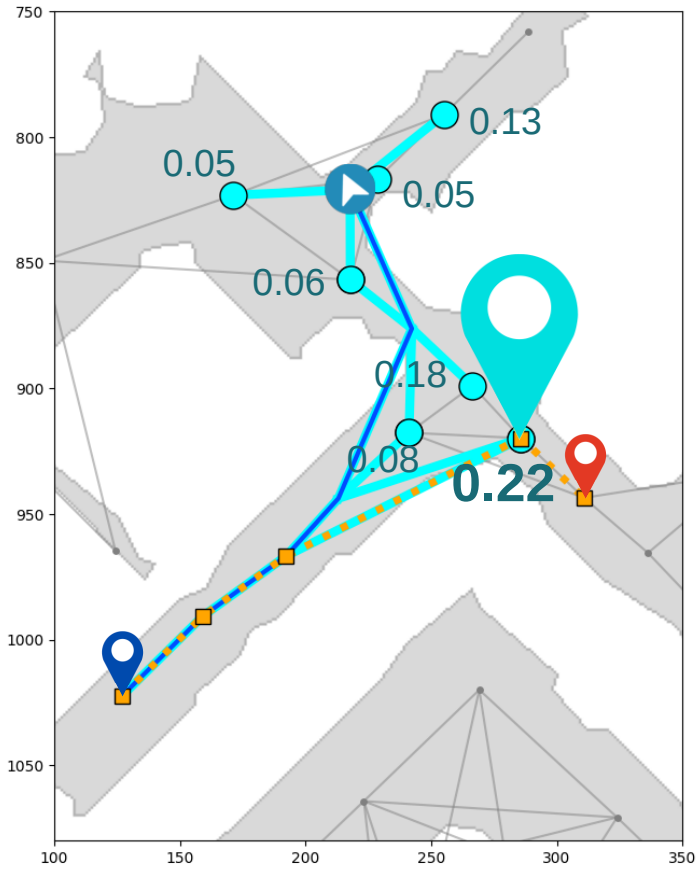
Unvisited & Observed Nodes

Experiment Visualization 2: Local Goal Search

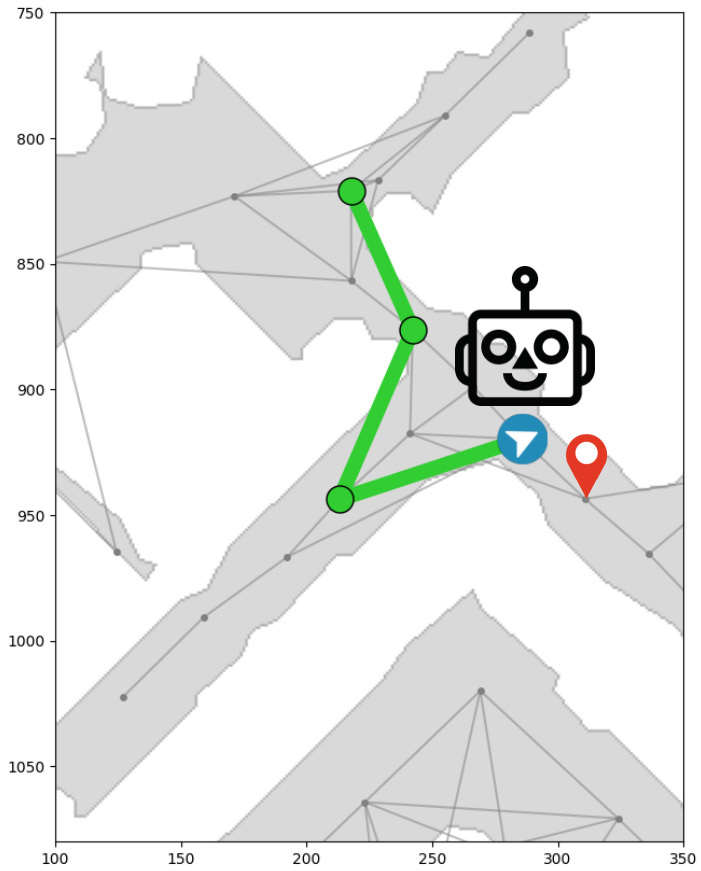
RLLAB
http://rllab.snu.ac.kr

Local goal search using the proposed **spectral-domain SOS features** significantly improves the **success rate** and **SPL**.

Local Goal Search



Exploit Trajectory



Ablation Study: Extension to Continuous Env.

RLLAB
http://rllab.snu.ac.kr

We extend **Meta-Explore** to **image-goal navigation task** in **continuous environments** to address the impact of **hierarchical exploration** in realistic environments.

navigation mode: exploite

Contributions

1) We propose a hierarchical navigation method called Meta-Explore, deploying an exploitation policy to correct misled recent actions. The agent searches for an appropriate local goal instead of reversing the recent action sequence.

2) In the exploitation mode, the agent uses a novel scene representation called scene object spectrum (SOS). SOS features provide semantically meaningful clues to choose a near-optimal local goal.

3) Meta-Explore shows better generalization results compared to all baselines in three VLN benchmarks.

RLLAB
http://rllab.snu.ac.kr

If you have any questions,

please contact to minyoung.hwang@rllab.snu.ac.kr

project page: https://rllab-snu.github.io/projects/Meta-Explore/doc.html

Thank you for your attention