The Regretful Agent: Heuristic-Aided Navigation through Progress Estimation

Chih-Yao Ma[1], Zuxuan Wu[2], Ghassan AlRegib[1], Caiming Xiong[3], and Zsolt Kira[1]

[1]Georgia Institute of Technology, [2]University of Maryland, College Park, [3]Salesforce Research

https://github.com/chihyaoma/selfmonitoring-agent
https://github.com/chihyaoma/regretful-agent

1 PROBLEM: VISION-AND-LANGUAGE NAVIGATION (VLN)

Exit the bedroom and go towards the table. Go to the stairs on the left of the couch. Wait on the third step.

Matterport3D

Panoramic camera

Challenges:

Association between language and images

Exit *Bedroom* →

No map or GPS

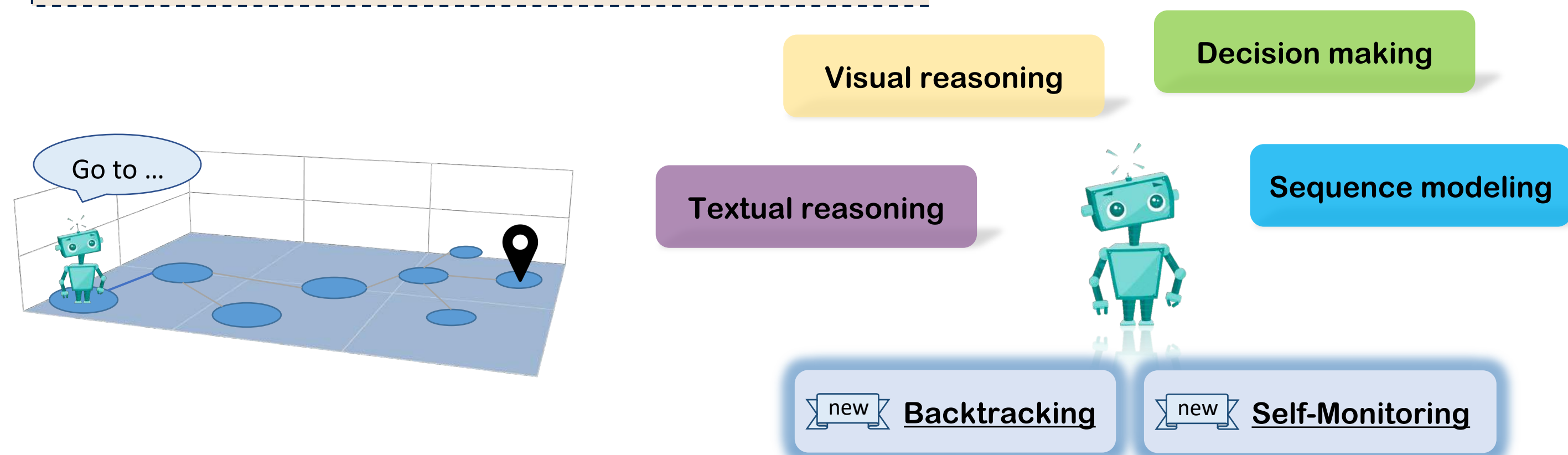**No explicit object as the goal**
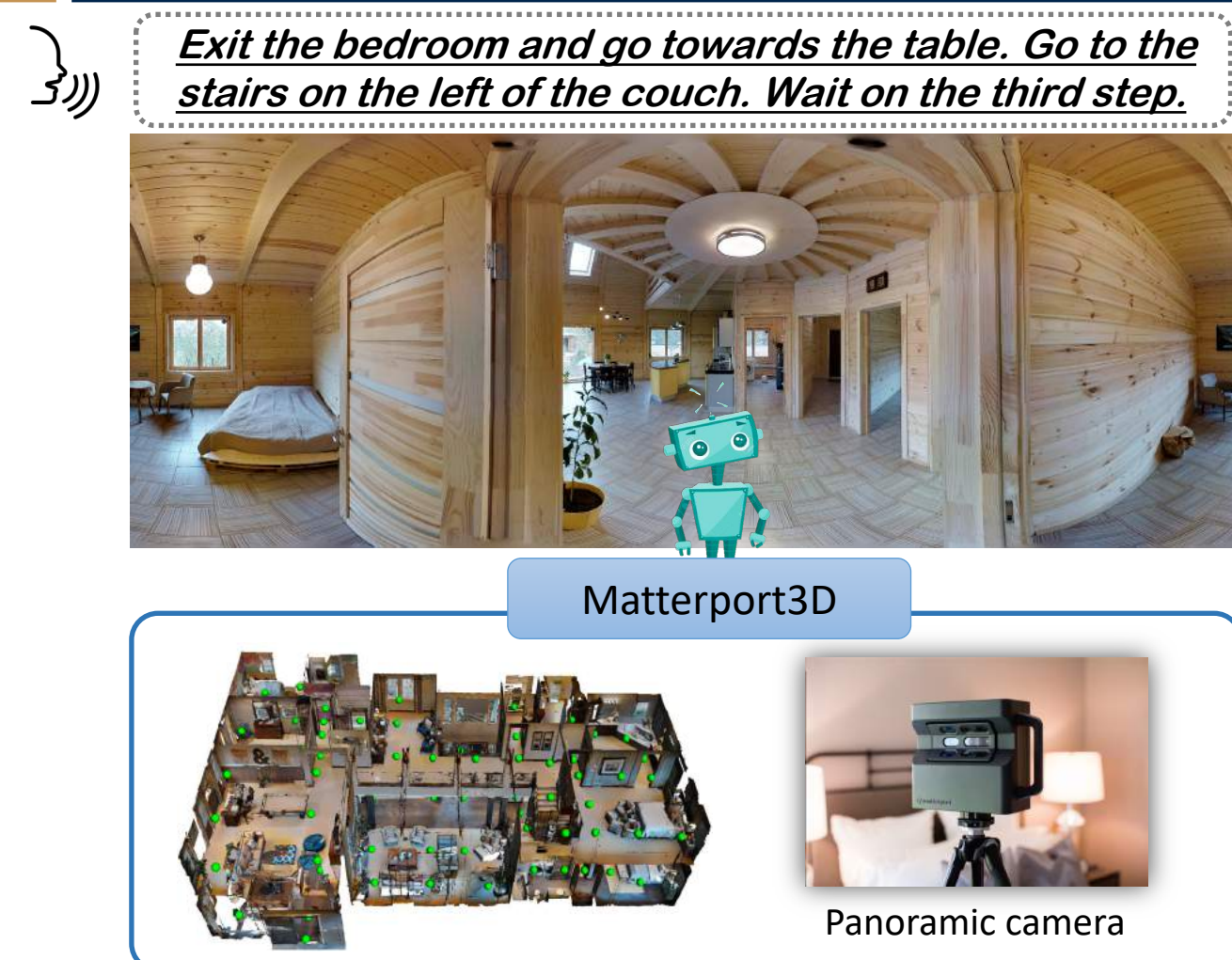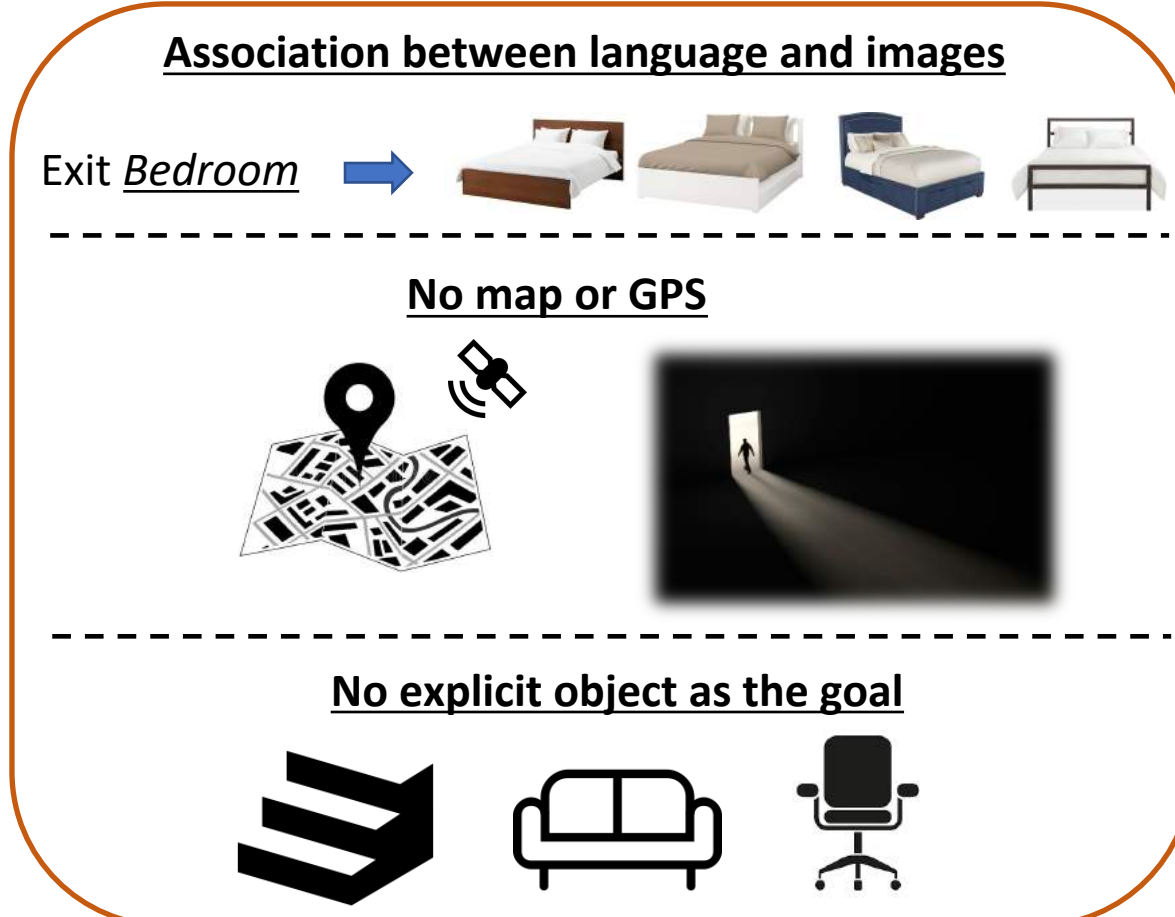
2 INTUITION: NAVIGATION AS GRAPH SEARCH

- Prior methods rely heavily on **beam search**, which requires a thorough exploration of trajectories. We instead focus on **greedy** navigation without frontiers.

- We view the problem as **graph search**, leveraging our prior work on a progress monitor – *a learned heuristic*.

- We propose to **learn a backtracking** mechanism in an end-to-end manner as well as **progress markers** to mark states.
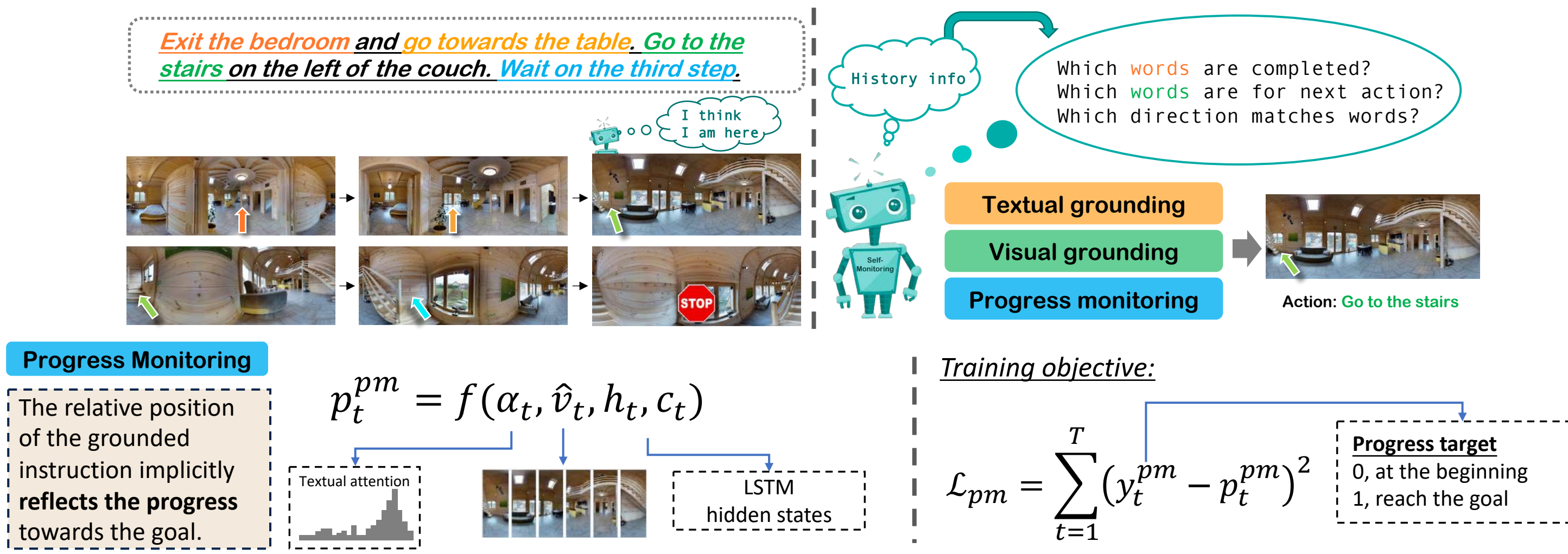
Greedy
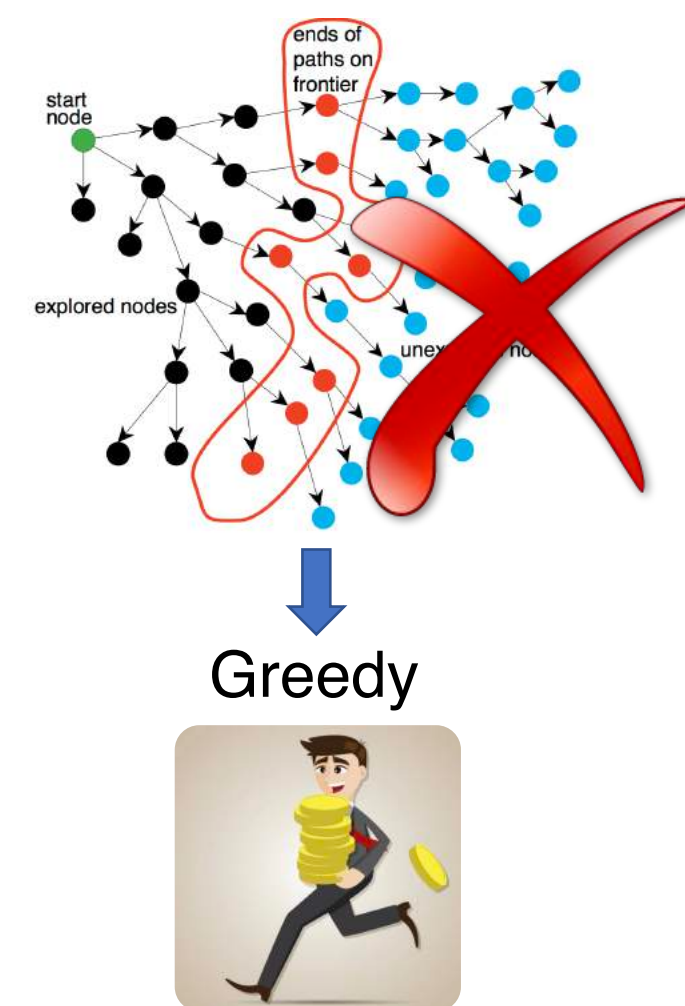
3 GOAL-DRIVEN REASONING WITH LEARNED BACKTRACKING
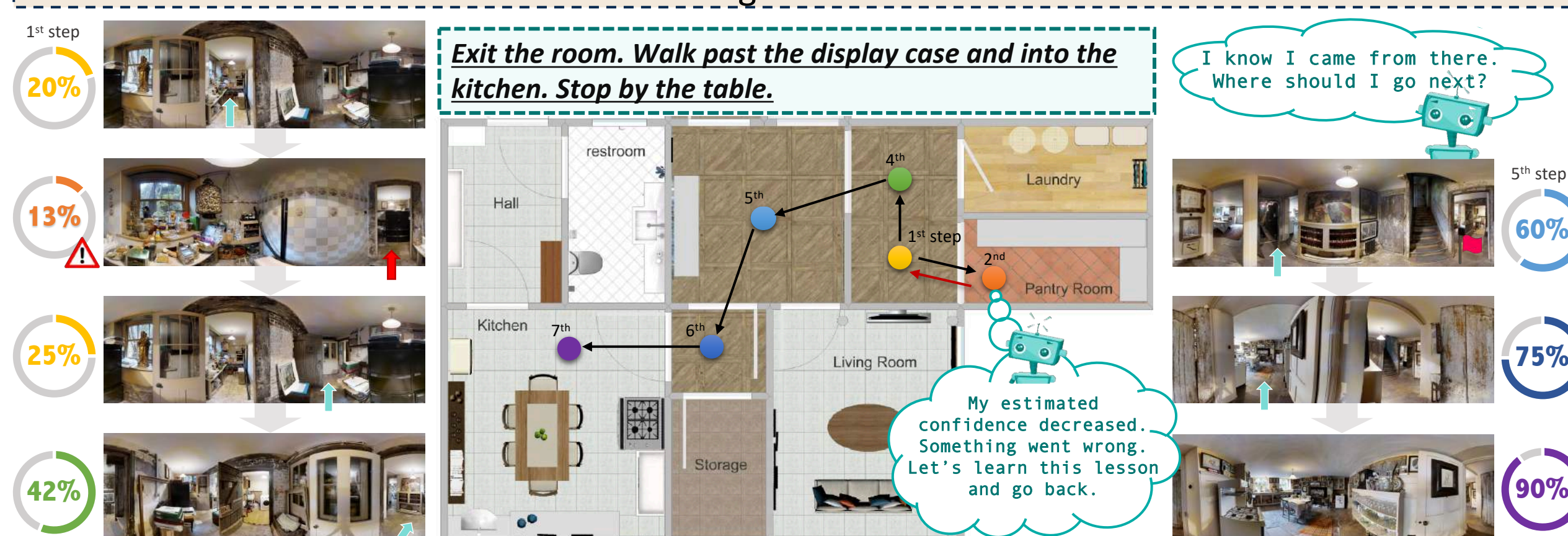
What do we need for goal-driven navigation?

Go to ...

Visual reasoning   Decision making

Textual reasoning   Sequence modeling

new **Backtracking**   new **Self-Monitoring**

4 SELF-MONITORING AGENT

ICLR

Exit the bedroom and go towards the table. Go to the stairs on the left of the couch. Wait on the third step.

History info

Which words are completed?
Which words are for next action?
Which direction matches words?

I think I am here

Textual grounding
Visual grounding
Progress monitoring

Action: Go to the stairs

Progress Monitoring

The relative position of the grounded instruction implicitly **reflects the progress** towards the goal.

$$p_t^{pm} = f(\alpha_t, \hat{v}_t, h_t, c_t)$$

Textual attention   LSTM hidden states

Training objective:

$$\mathcal{L}_{pm} = \sum_{t=1}^{T} (y_t^{pm} - p_t^{pm})^2$$

Progress target
0, at the beginning
1, reach the goal

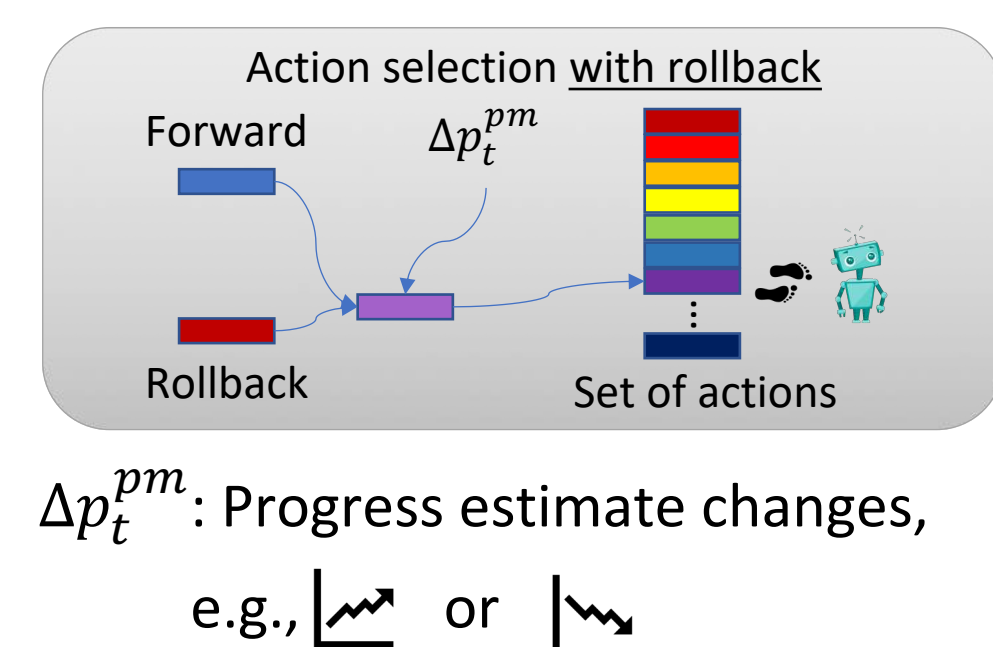## 1 END-TO-END LEARNED BACKTRACKING AGENT

CVPR

Leverages the self-monitoring mechanism through time to decide when to **roll back** to a previous location and resume the instruction following task.
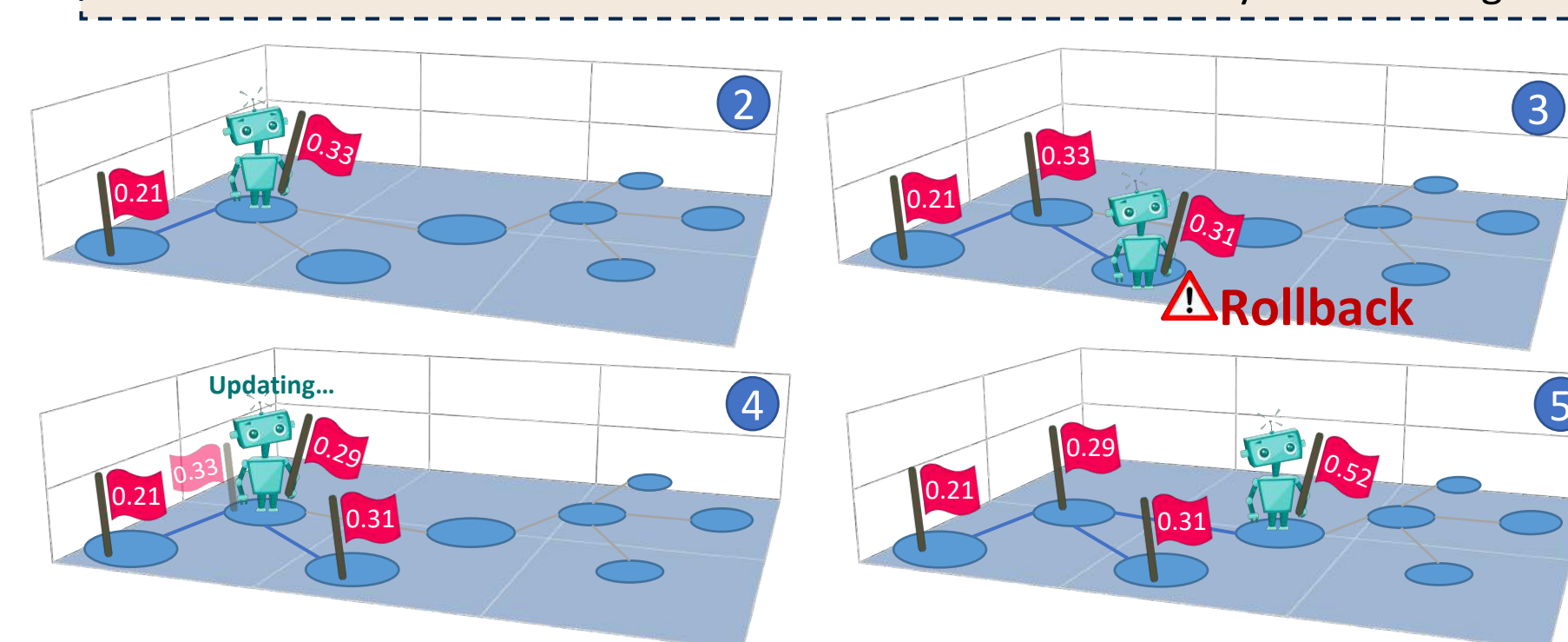
1st step
20%
13%
25%
42%

Exit the room. Walk past the display case and into the kitchen. Stop by the table.

I know I came from there. Where should I go next?

5th step
60%
75%
90%

My estimated confidence decreased. Something went wrong. Let's learn this lesson and go back.

2 BACKTRACKING FRAMEWORK

Regret Module: forward or rollback?

Action selection with rollback

Forward Δp_t^{pm}

Rollback   Set of actions

Δp_t^{pm}: Progress estimate changes,
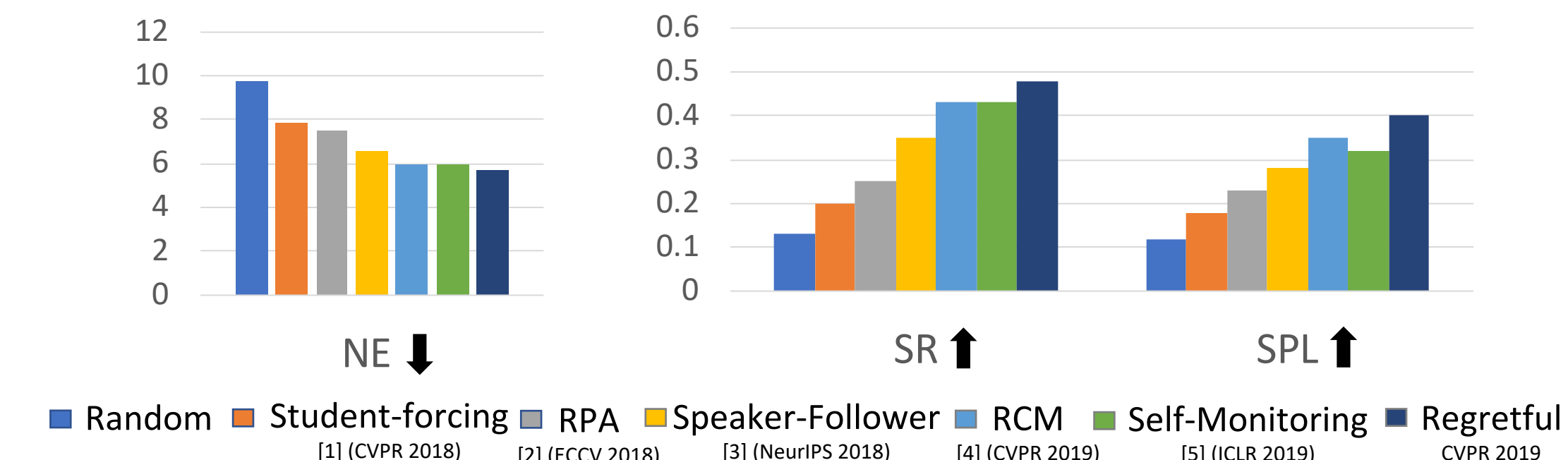
e.g.,   or

Progress Marker: which way to go?

Local Graph Search:
- know which directions have been visited
- estimate which one is likely lead to the goal.

Updating...

⚠️Rollback

3 QUANTITATIVE ANALYSIS

NE ↓   SR ↑   SPL ↑

■ Random ■ Student-forcing ■ RPA ■ Speaker-Follower ■ RCM ■ Self-Monitoring ■ Regretful
[1] (CVPR 2018) [2] (ECCV 2018) [3] (NeurIPS 2018) [4] (CVPR 2019) [5] (ICLR 2019) CVPR 2019
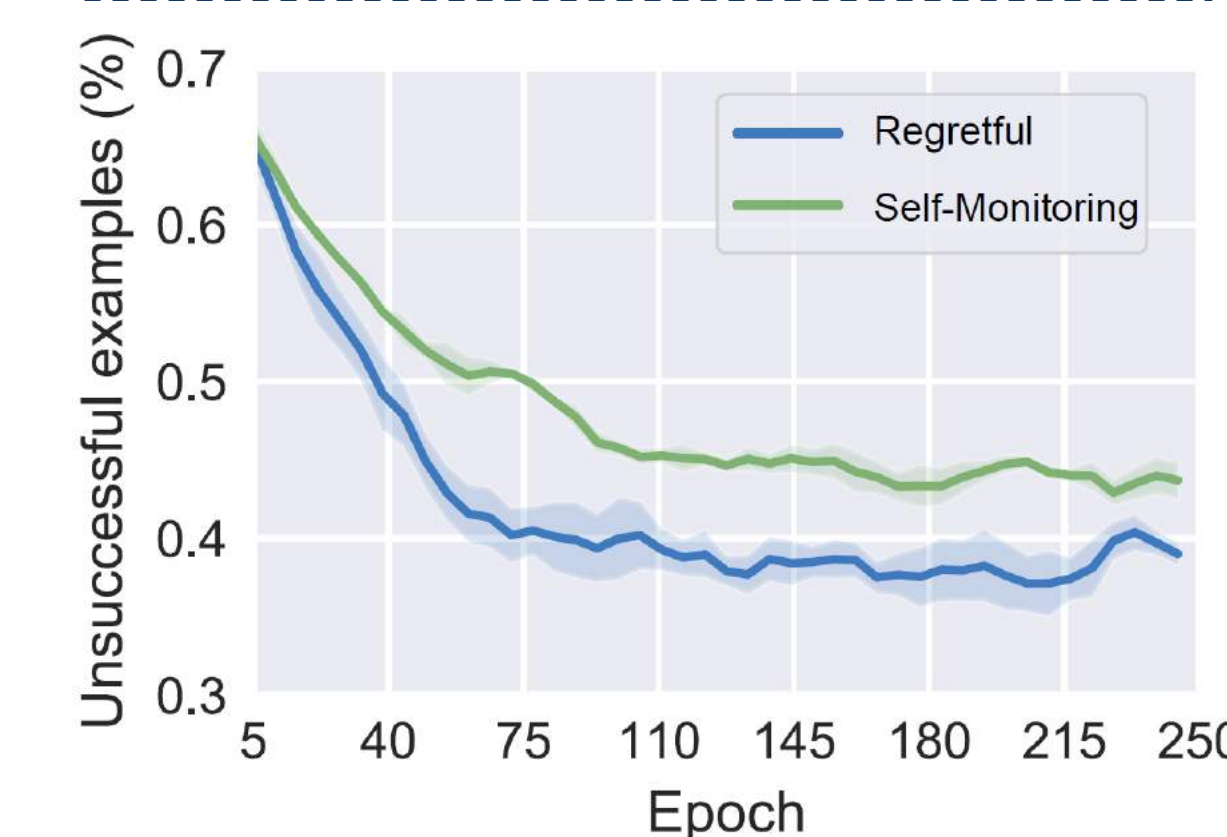
Evaluation metrics:
- **Navigation Error (NE)**: mean of the shortest path distance between the agent's final position and the goal location.
- **Success Rate (SR)**: the percentage of final positions less than 3m away from the goal location.
- **Success rate weighted by Path Length (SPL)**: SR weighted with trajectory lengths.

4 QUALITATIVE RESULTS & OTHER ANALYSIS

Walk down the hall way and make a right at the stairs and walk down the stairs. Make a hard left at the bottom of the stairs and wait by the Bamboo plant.

1 Steps: 1 / 8 Progress Monitor: 0.16 **Forward**   16%
2 Steps: 2 / 8 Progress Monitor: 0.19 ⚠️ **Rollback**   19%
3 Steps: 3 / 8 Progress Monitor: 0.21 **Forward**   21%
8 Steps: 8 / 8 Progress Monitor: 0.99 STOP Stop   99%

Number of unsuccessful examples **involving rollback action** reduced

— Regretful
— Self-Monitoring

Sanity check: manually blocking rollback

Val-Unseen	Rollback	NE	SR
Regretful	🚫	5.80	0.46
Regretful	✓	**5.36**	**0.48**

Reference:
[1] Anderson et al., "Vision-and-Language Navigation: Interpreting visually-grounded navigation instructions in real environments", CVPR 2018.
[2] Wang et al., "Look Before You Leap: Bridging Model-Free and Model-Based Reinforcement Learning for Planned-Ahead Vision-and-Language Navigation", ECCV 2018.
[3] Fried et al., "Speaker-Follower Models for Vision-and-Language Navigation", Neural 2018.
[4] Wang et al., "Reinforced Cross-Modal Matching and Self-Supervised Imitation Learning for Vision-Language Navigation", CVPR 2019.
[5] Ma et al., "Self-Monitoring Navigation Agent via Auxiliary Progress Estimation", ICLR 2019.