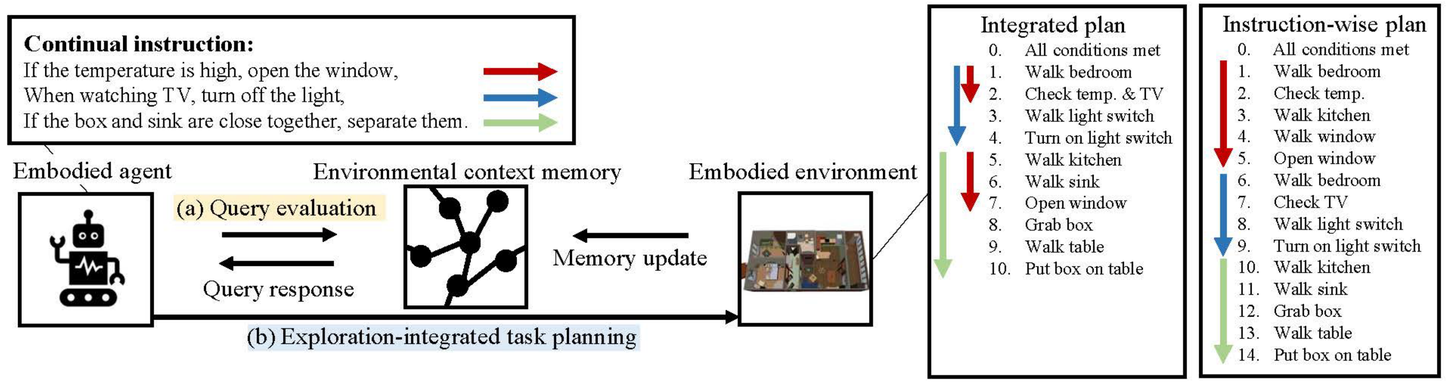
Exploratory Retrieval-Augmented Planning For Continual Embodied Instruction Following

Minjong Yoo[1], Jinwoo Jang[1], Wei-jin Park[2], Honguk woo[1*]

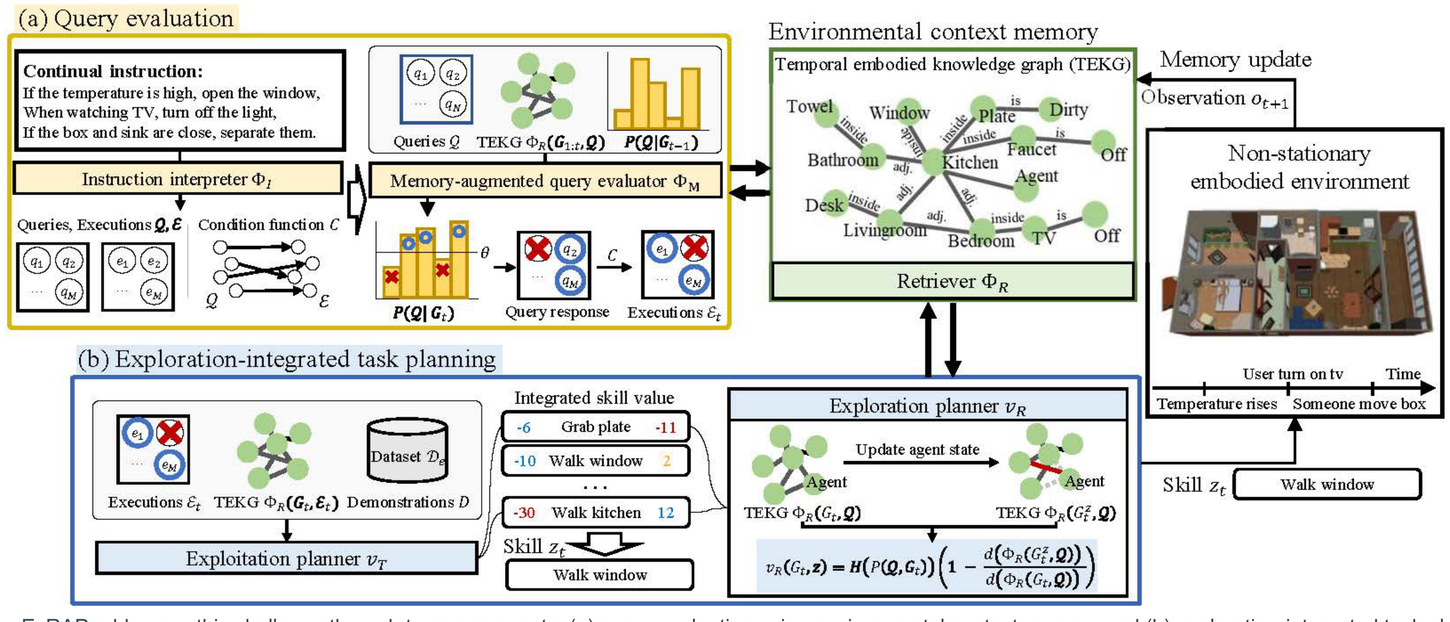
[1]Department of computer science and engineering, Sungkyunkwan University [2]Acryl Inc.

Continual instruction following



- We consider a set of instructions for embodied tasks that are continuously and simultaneously conducted based on specific environmental contexts.
- Embodied tasks are conditioned on environmental contexts and conducted **continuously**, as **continual instruction following**, aligning closely with continual queries that monitor updates of interest and return results when specific thresholds are met.
- We propose a novel **ExRAP framework**, systematically combining LLMs' reasoning capabilities and environmental context memory into exploration-integrated task planning to tackle continuous instruction following tasks in non-stationary embodied environments.

Exploratory Retrieval-Augmented Planning (ExRAP)

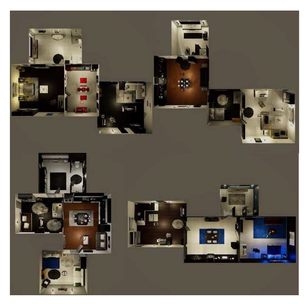


- ExRAP addresses this challenge through two components: (a) query evaluation using environmental context memory and (b) exploration-integrated task planning.
- In (a), a temporal embodied knowledge graph creates the environmental context memory, enabling an LLM-based query evaluator to assess conditions and confidence, refined by information-based temporal consistency to handle synchronization uncertainties.
- In (b), ExRAP plans skills that balance task achievement (**exploitation**) and confidence in query evaluations (**exploration**) by integrating exploitation values from LLM in-context learning with exploration values based on information estimation.

Experiments

- Test Environment: ALFRED, VirtualHome, CARLA

VirtualHome



ALFRED

CARLA

Model	Low non-stationarity		Medium non-stationarity		High non-stationarity	
	SR (↑)	PS (↓)	SR (↑)	PS (↓)	SR (↑)	PS (↓)
Evaluation in VirtualHome						
ZSP	20.59%±4.71%	31.03±4.68	20.06%±1.93%	32.06±4.66	17.28%±3.16%	24.08±4.63
SayCan	35.12%±4.83%	21.67±3.81	33.69%±5.36%	21.81±4.14	27.33%±4.24%	16.18±3.98
ProgPrompt	32.10%±4.41%	18.84±4.08	30.51%±5.31%	23.43±1.07	27.19%±2.99%	18.60±4.22
LLM-Planner	40.97%±7.00%	17.61±1.40	39.89%±4.52%	15.93±2.13	34.60%±6.49%	14.94±2.89
ExRAP	**61.12%**±7.03%	**11.75**±2.49	**55.14%**±6.59%	**11.33**±1.92	**50.12%**±5.70%	**8.61**±2.25
Evaluation in ALFRED						
ZSP	18.22%±5.33%	17.24±2.12	14.67%±6.18%	20.83±3.63	9.56%±4.80%	22.53±3.57
SayCan	45.67%±6.89%	8.25±1.86	41.81%±7.64%	8.39±3.55	35.79%±6.31%	7.42±1.14
ProgPrompt	47.15%±1.17%	9.81±2.14	35.62%±1.04%	7.22±1.35	19.97%±0.80%	7.52±2.45
LLM-Planner	58.44%±3.97%	7.28±1.09	51.80%±3.79%	7.28±1.05	35.76%±6.00%	6.65±1.06
ExRAP	**69.90%**±1.47%	**5.94**±0.92	**64.00%**±5.07%	**4.82**±1.03	**59.11%**±2.48%	**4.42**±1.36
Evaluation in CALRA						
ZSP	10.44%±1.03%	29.35±7.21	6.89%±2.98%	32.46±6.03	4.67%±1.17%	33.00±1.40
SayCan	37.55%±4.74%	20.73±5.36	35.11%±6.12%	22.44±5.38	30.71%±5.28%	21.71±2.01
LLM-Planner	50.83%±1.60%	14.02±3.01	44.00%±0.70%	14.39±1.94	41.58%±3.35%	13.59±2.65
ExRAP	**65.25%**±7.47%	**12.43**±3.90	**62.25%**±6.72%	**11.50**±2.24	**58.83%**±10.08%	**10.84**±2.52

- Task Success Rate (**SR**): Measures the proportion of tasks completed for continual instructions with conditions met at each timestep.
- Pending Step (**PS**): Represents the average timesteps required to complete tasks once the conditions of the instructions are satisfied in the environment.
- ExRAP consistently outperforms other models. in terms of SR across different non-stationarity levels and environments
- ExRAP achieves lower PS compared to other models, particularly in medium and high non-stationarity scenarios, reflecting its efficiency in task completion even under challenging conditions.