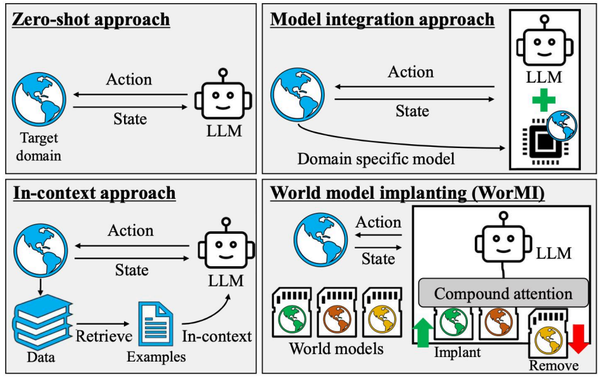
World Model Implanting for Test-time Adaptation of Embodied Agents

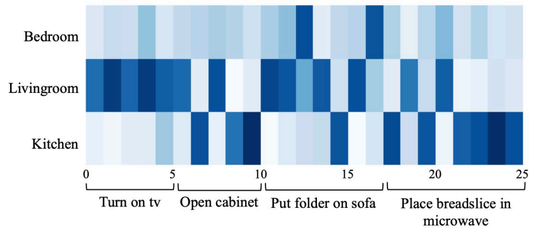
Minjong Yoo • Jinwoo Jang • Sihyung Yoon • Honguk Woo*

Contributions



- We present the novel **WorMI** framework to enable cross-domain embodied policy adaptation at test-time, exploring the dual-stage model compositionality: **world-to-world knowledge integration and world-to-reasoning alignment**.
- This dual-stage design not only combines the strengths of model integration (which embeds a domain-specific model as part of the policy) and in-context adaptation (which incorporates examples relevant to target domains) but also significantly enhances adaptability to unseen domains.

Visualization of world-level attention map



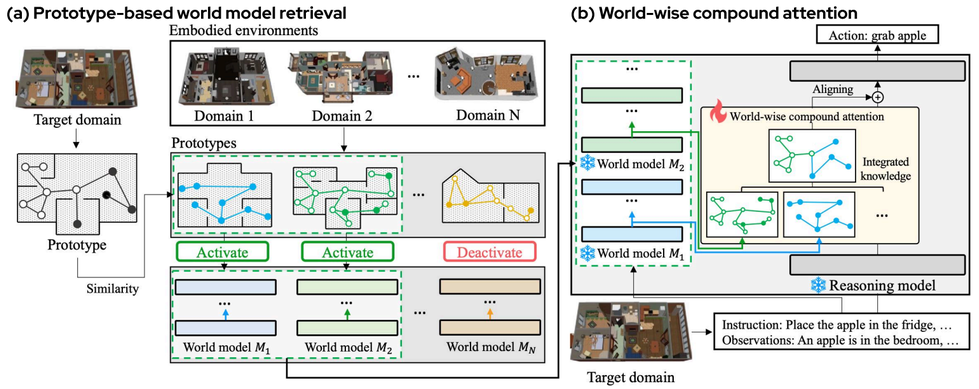
- WorMI dynamically shifts its focus among the three world models as the target domain varies along with different tasks, assigning higher attention weights to the model most relevant to the current task.

Our Approach

WorMI Framework

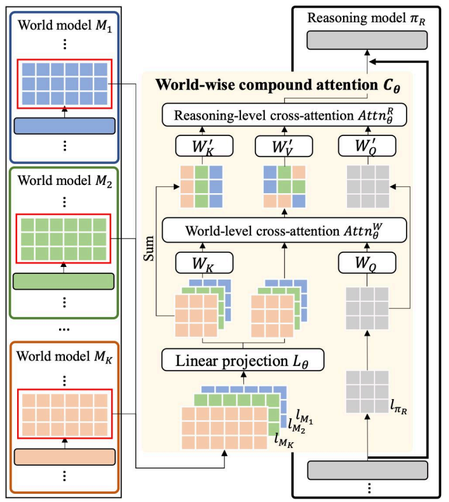
(a) Prototype-based world model retrieval: Retrieves and integrates only the most relevant world models for the current target domain

(b) World-wise compound attention: Integrates the intermediate representations of multiple world models and then align them with the reasoning model



World-wise Compound Attention

- **Linear Projection:** matching its dimension to that of the reasoning model's embedding space
- **World-level Cross-attention:** quantifying the relative importance of each world model's representation, conditioned on the current hidden state of the reasoning model.
- **Reasoning-level Cross-attention:** after obtaining the integrated representation of world models via the world-level cross-attention, we align it to the reasoning model.



Evaluations

- Evaluated in VirtualHome and ALFRED environments.

VirtualHome

ALFRED

- **Cross-domain environment:** Various room layouts (Scenes) and tasks types (Tasks)
- **Performance metric: SR** is the Success Rate, and **PS** is the Pending Steps.

Zero-shot results

Model	Seen Tasks & Seen Scenes		Seen Tasks & Unseen Scenes		Unseen Tasks & Unseen Scenes	
	SR (↑)	PS (↓)	SR (↑)	PS (↓)	SR (↑)	PS (↓)
Evaluation in VirtualHome						
ZSP	11.15%±0.65%	29.02±0.08	8.95%±1.13%	29.25±0.11	8.19%±0.33%	29.36±0.56
LLM-FT	58.55%±2.18%	17.37±0.34	53.42%±0.79%	17.70±0.23	42.82%±0.76%	20.79±0.17
LLM-Planner	35.67%±1.25%	27.15±0.13	28.55%±0.42%	27.04±0.12	21.45%±0.42%	27.73±0.05
SayCanPay	69.88%±2.32%	14.53±0.55	64.74%±0.87%	15.64±0.18	45.71%±0.59%	19.04±0.63
WorMI	**85.78%**±0.45%	**10.76**±0.19	**80.26%**±1.02%	**12.42**±0.09	**66.12%**±0.80%	**15.17**±0.08
Evaluation in ALFWorld						
ZSP	2.30%±0.21%	49.34±0.66	2.26%±0.09%	49.04±0.95	2.13%±0.09%	49.68±0.23
LLM+FT	45.72%±0.39%	14.63±1.35	29.26%±0.72%	38.49±2.87	29.40%±1.65%	46.84±0.15
LLM-Planner	17.73%±0.61%	32.09±2.90	18.63%±0.82%	40.50±2.57	12.31%±0.80%	46.33±1.68
SayCanPay	40.67%±1.24%	18.37±1.93	34.10%±1.17%	20.82±1.21	39.66%±1.43%	23.86±1.90
WorMI	**62.51%**±1.65%	**9.96**±1.29	**52.67%**±1.39%	**17.74**±0.83	**51.67%**±2.23%	**20.18**±0.63

Few-shot results

Model	1-Shot		5-Shot		10-Shot	
	SR (↑)	PS (↓)	SR (↑)	PS (↓)	SR (↑)	PS (↓)
Evaluation in VirtualHome						
LLM-FT	42.35%±0.85%	20.46±0.15	47.22%±0.46%	16.82±0.29	51.45%±1.06%	16.82±0.39
LLM-Planner	24.63%±0.34%	26.19±0.31	29.80%±0.55%	26.98±0.08	33.49%±0.44%	26.60±0.16
SayCanPay	46.75%±1.02%	19.12±0.16	49.80%±1.11%	15.52±0.11	56.24%±0.79%	16.00±0.13
WorMI	**74.90**±1.57%	**12.18**±0.31	**78.04%**±0.34%	**11.85**±0.08	**79.61%**±0.99%	**11.68**±0.13
Evaluation in ALFWorld						
LLM-FT	26.78%±1.20%	46.51±0.27	29.79%±1.36%	46.21±0.37	33.57%±0.98%	43.70±0.83
LLM-Planner	18.28%±1.06%	43.94±2.87	16.80%±1.46%	40.50±3.14	17.76%±1.25%	42.45±2.64
SayCanPay	40.94%±1.34%	21.58±1.14	36.70%±1.31%	22.15±3.55	39.54%±1.48%	24.74±3.48
WorMI	**51.50%**±2.21%	**21.46**±3.96	**58.69%**±1.94%	**13.14**±1.00	**64.46%**±0.88%	**11.79**±0.81

- We evaluate the embodied task planning performance, wherein each policy is evaluated in a zero-shot and few-shot manner.
- WorMI outperforms LLM-based baselines with 16.2% higher SR and 3.77 lower PS in a zero-shot, and 22.87% higher SR and 6.17 lower PS in a few-shot manner.