

# From Assumptions to Actions: Turning LLM Reasoning into Uncertainty-Aware Planning for Embodied Agents

SeungWon Seo\*, SooBin Lim\*, SeongRae Noh\*,  
Haneul Kim, HyeongYeop Kang†

Korea University, Seongbuk-gu, Republic of Korea



**ICLR**  
International Conference On  
Learning Representations



**KOREA**  
UNIVERSITY

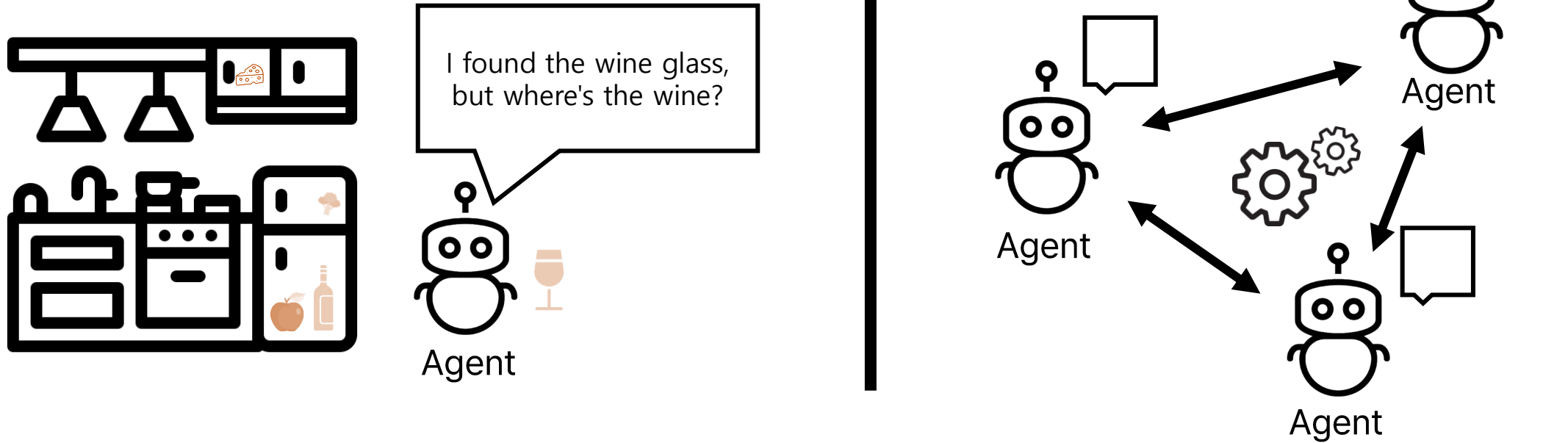


III XR Lab

\* These authors contributed equally.

† Corresponding author.

# Challenge



- Agents operate in decentralized and partially observable environments, relying only on local observations and limited communication.
- Prior approaches address uncertainty by relying heavily on frequent inter-agent communication.
- This leads to high LLM costs and can disrupt efficient collaboration, especially when humans are involved.

# Key Observation

## (a) Example of Planner reasoning trace

[goexplore] : action      could potentially lead you to one of the items you need : assumption

Okay, let's analyze the situation. My goal is to put items onto the <coffeetable>. I am not currently holding anything, and I need to move the following items: 2 <cupcakes>, 1 <apple>, 1 <wine>, and 1 <pudding>.

Let's consider the available actions and our goal to determine the optimal action.

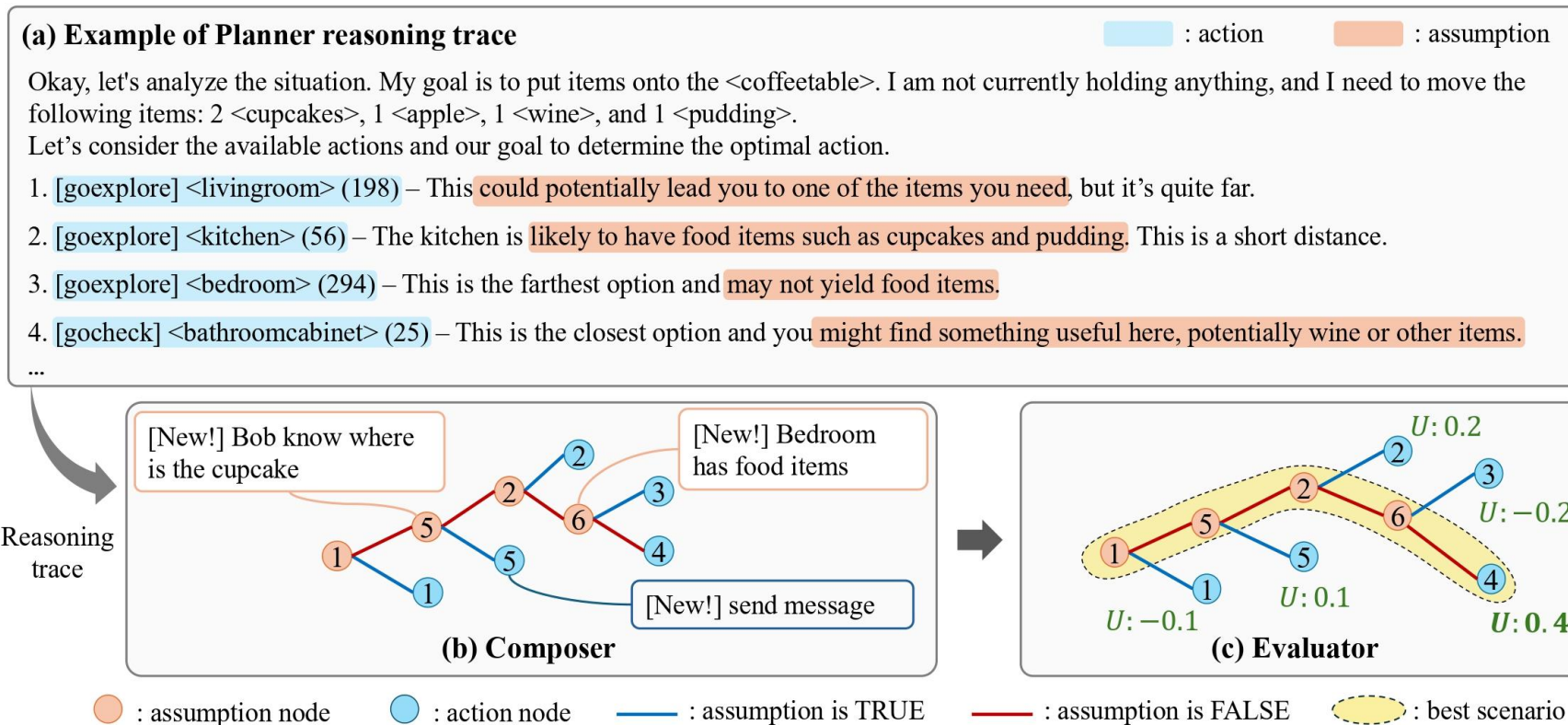
1. [goexplore] <livingroom> (198) – This could potentially lead you to one of the items you need, but it's quite far.
2. [goexplore] <kitchen> (56) – The kitchen is likely to have food items such as cupcakes and pudding. This is a short distance.
3. [goexplore] <bedroom> (294) – This is the farthest option and may not yield food items.
4. [gocheck] <bathroomcabinet> (25) – This is the closest option and you might find something useful here, potentially wine or other items.

...



- LLM planners implicitly generate local assumptions in their reasoning traces, but these remain fragmented and unstructured.
- As a result, the model cannot systematically compare the plausibility of assumptions or the utility of actions under each case.
- We instead construct a scenario tree by explicitly branching over uncertain environmental assumptions.

# PCE Framework



- PCE operates in three stages: the Planner generates an initial action and reasoning trace, and the Composer converts it into a scenario tree over assumptions.
- The Evaluator scores each scenario based on likelihood, conditional gain, and execution cost to select the optimal action.
- Communication is treated as one action among others, enabling explicit decisions between communicating and acting under uncertainty.



# Experiment Environment



**C-WAH Environment**



**TDW-MAT Environment**

- We evaluate our method on two long-horizon, multi-room household benchmarks: C-WAH and TDW-MAT.
- These environments require agents to operate under partial observability across multiple rooms over extended time horizons.
- We test across diverse LLM backbones, including GPT-4o mini, GPT-OSS:20B, and Gemma3:4B.

# Experiment Environment

		PCE	CoELA	REVECA	CaPo	CoTS
GPT-4o mini	<i>Total Steps</i> ↓	<b>42.76</b>	60.40	<u>46.80</u>	60.82	64.00
	<i>Comm</i>	1.70	9.88	6.00	8.72	10.24
	<i>Usages</i> ↓	<u>44353.56</u>	55467.12	46312.16	<b>41702.00</b>	44628.12
GPT-OSS:20B	<i>Total Steps</i> ↓	<b>49.60</b>	72.72	<u>53.86</u>	68.34	65.26
	<i>Comm</i>	2.11	9.22	6.49	8.20	8.32
	<i>Usages</i> ↓	<b>73535.24</b>	77727.20	<u>74764.24</u>	99810.44	95428.84
Gemma3:4B	<i>Total Steps</i> ↓	<b>59.20</b>	77.20	<u>62.56</u>	75.88	72.32
	<i>Comm</i>	3.02	9.48	9.14	7.92	4.04
	<i>Usages</i> ↓	50984.7	<u>49271.24</u>	<b>44637.58</b>	64015.24	51966.64

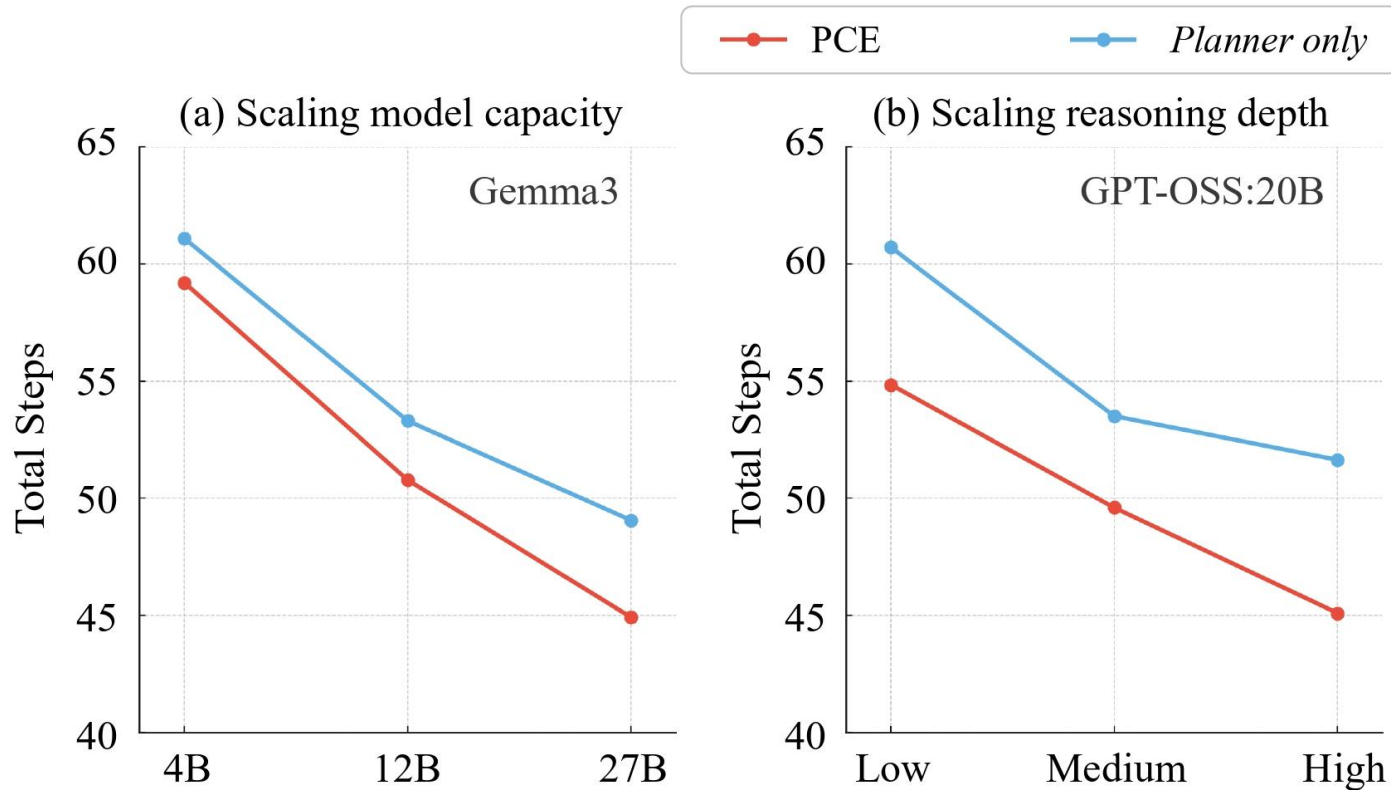
		PCE	CoELA	REVECA	CaPo	CoTS
GPT-4o mini	<i>Total</i> ↑	<b>87.50</b>	62.50	<u>81.25</u>	73.33	75.00
	<i>Food</i> ↑	<b>89.17</b>	65.83	<u>80.83</u>	82.50	84.17
	<i>Stuff</i> ↑	<b>85.83</b>	59.17	<u>81.66</u>	64.17	65.83
	<i>Comm</i>	3.58	13.33	43.76	70.79	108.92
	<i>Usages</i> ↓	197807.29	<b>113058.83</b>	<u>185453.54</u>	281404.71	411392.08
GPT-OSS:20B	<i>Total</i> ↑	<b>81.25</b>	55.00	<u>73.33</u>	65.41	59.17
	<i>Food</i> ↑	<b>85.00</b>	50.83	<u>78.33</u>	72.50	70.83
	<i>Stuff</i> ↑	<b>77.50</b>	59.17	<u>68.33</u>	58.33	47.50
	<i>Comm</i>	13.75	11.62	107.79	43.00	41.83
	<i>Usages</i> ↓	337225.12	<b>237498.88</b>	370737.17	348066.92	<u>334912.67</u>
Gemma3:4B	<i>Total</i> ↑	<b>70.83</b>	45.84	52.09	<u>67.50</u>	63.33
	<i>Food</i> ↑	<b>71.66</b>	50.00	56.67	<u>70.83</u>	64.17
	<i>Stuff</i> ↑	<b>70.00</b>	41.67	47.50	<u>64.17</u>	62.50
	<i>Comm</i>	9.42	27.42	108.00	57.88	55.96
	<i>Usages</i> ↓	<u>184809.08</u>	<b>98350.25</b>	308221.25	217626.50	212029.79

## C-WAH

- PCE achieves the fastest goal completion in C-WAH and the highest success rates in TDW-MAT across all tested backbones.
- These gains come from explicitly structuring uncertainty, enabling agents to reduce unnecessary communication and focus on productive actions.

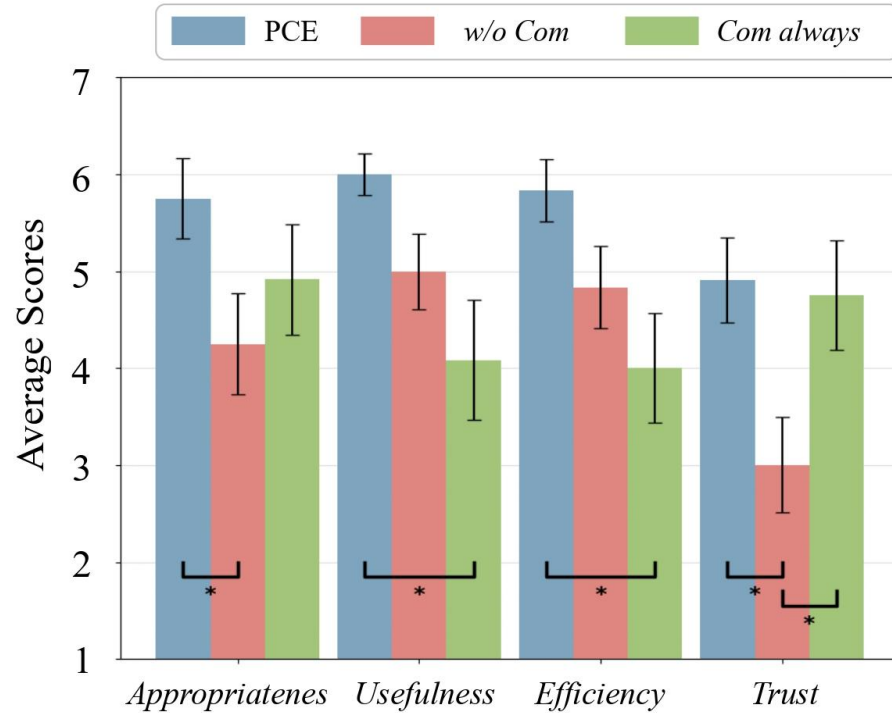
## TDW-MAT

# Experiment Environment



- Scaling model capacity or reasoning depth yields only modest improvements in the Planner-only baseline.
- This indicates that fragmented and implicitly handled assumptions persist even under increased model scaling.
- In contrast, PCE explicitly structures and evaluates assumptions, achieving consistent performance gains across backbones.

# Experiment Environment



## Appropriatenes

Did the agent perform actions appropriate to your intentions?

## Usefulness

Did the agent's performance contribute to achieving the goal efficiently?

## Efficiency

Did the agent's performance contribute to achieving the goal efficiently?

## Trust

Did you feel a sense of trust with the agent?

- A user study with 12 participants in C-WAH evaluates human-agent collaboration using a 7-point Likert scale.
- PCE achieves the highest scores across all criteria, including appropriatenes, usefulness, efficiency, and trust.
- This shows that effective collaboration comes from selectively triggering communication only when it is truly beneficial.

# Conclusion

- Explicitly structuring implicit assumptions in LLM reasoning into a scenario tree.
- Enabling uncertainty-aware action selection with a Planner–Composer–Evaluator framework.
- Showing consistent gains across benchmarks and backbones beyond communication-heavy strategies or mere scaling.