In-situ Value-aligned Human-Robot Interactions with Physical Constraints

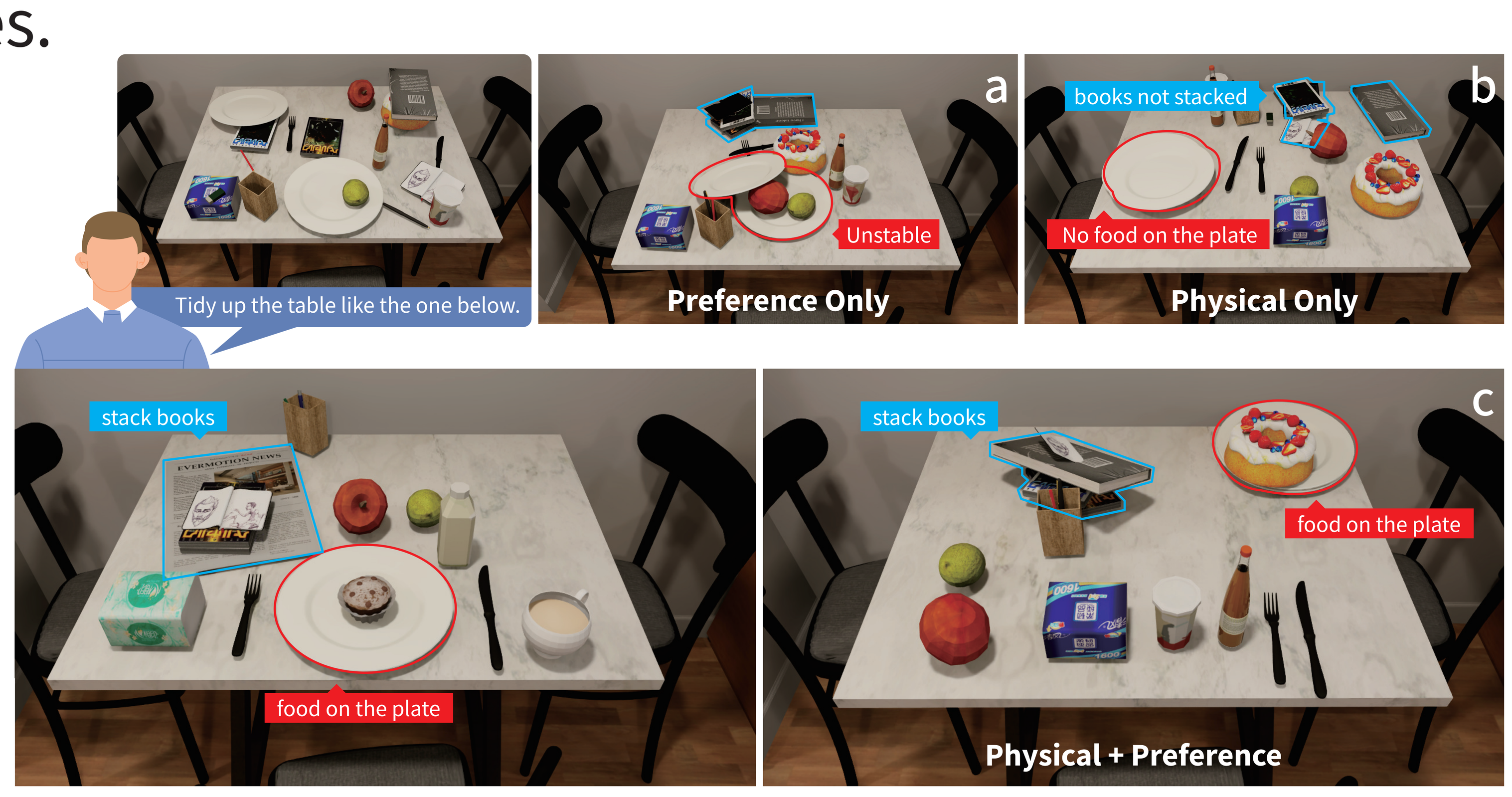
Hongtao Li[1,2], Ziyuan Jiao[2], Xiaofeng Liu[1,*], Hangxin Liu[2,*], Zilong Zheng[2,*]

[1]College of Artificial Intelligence and Automation, Hohai University

[2]State Key Laboratory of General Artificial Intelligence, BIGAI, Beijing, China

Motivation

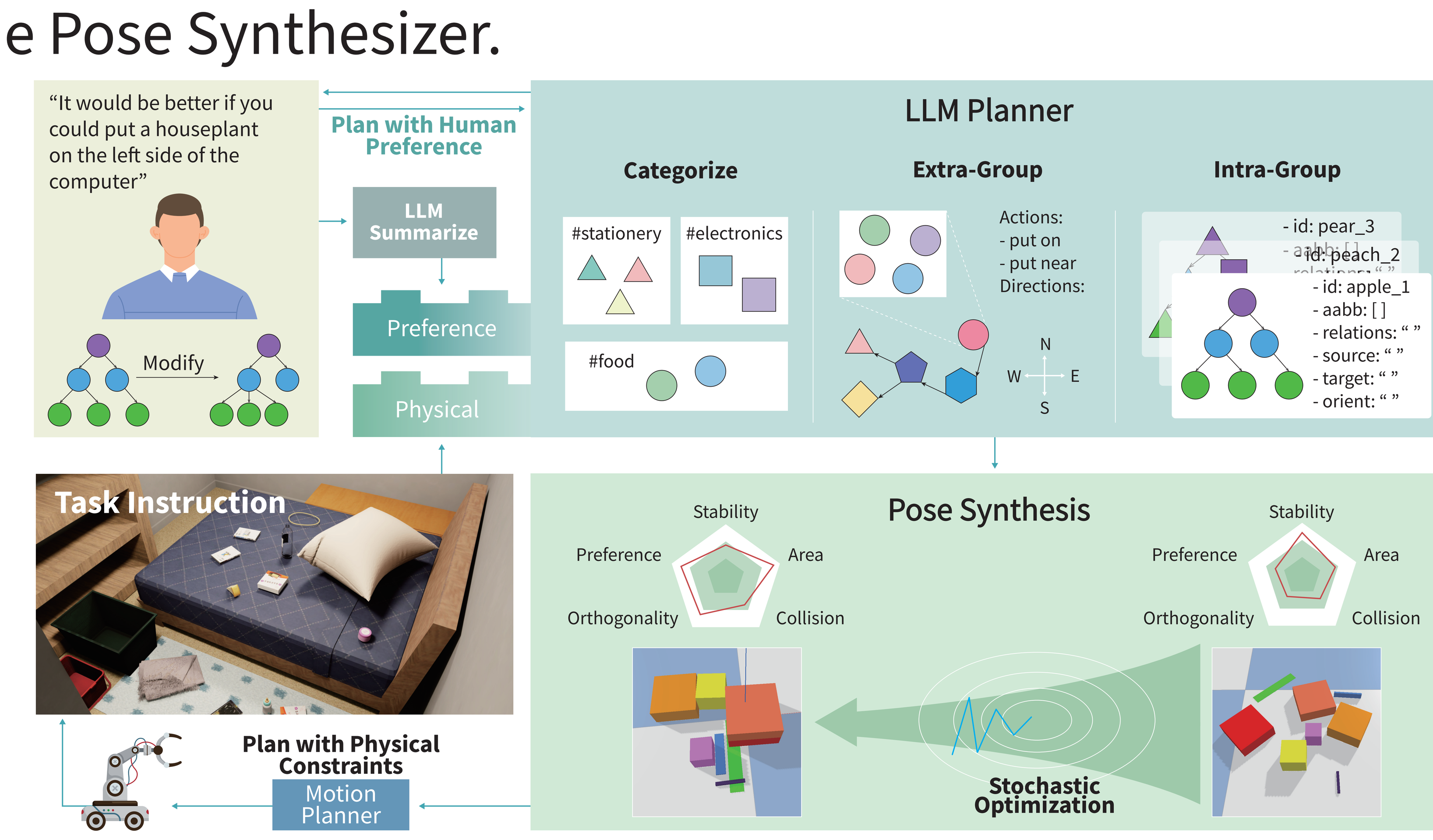
Imagine a scenario where robots tidy up a table, and humans expect the robots to do so according to their preferences.



(a). Human preferences only → unrealistic behavior.
(b). Physical constraints only → fail to meet the human expectation.
(c). Physical constraints + Human Preference → complete the task satisfactorily.

Methodology

We propose In-Context Learning from Human Feedback (ICLHF), which is capable of **learning human preferences in situ** and **combining them with physical constraints** to accomplish tasks. It primarily includes the LLM Planner and the Pose Synthesizer.



- The LLM Planner generates object placement plans using in-context learning based on preferences and constraints.
- We integrated a customized version of POG, an algorithm for efficient sequential manipulation planning on scene graphs, to enhance the plans generated by the LLM planner. Additional objective functions:

$$\mathcal{L}_{\text{manhattan}} := \sum_{l \in \mathcal{G}} \mathbf{1}_{|l|>1} \sum_{\mathbf{m},\mathbf{n} \in l} \|\mathbf{m}-\mathbf{n}\|_1$$

→ constrains the distance between every two objects

$$\mathcal{L}_{\text{area}} := \mathcal{L}_{\text{manhattan}} + \sum_{l \in \mathcal{G}} \mathbf{1}_{|l|>1} R(\mathbf{x}^l) \cdot R(\mathbf{y}^l)$$

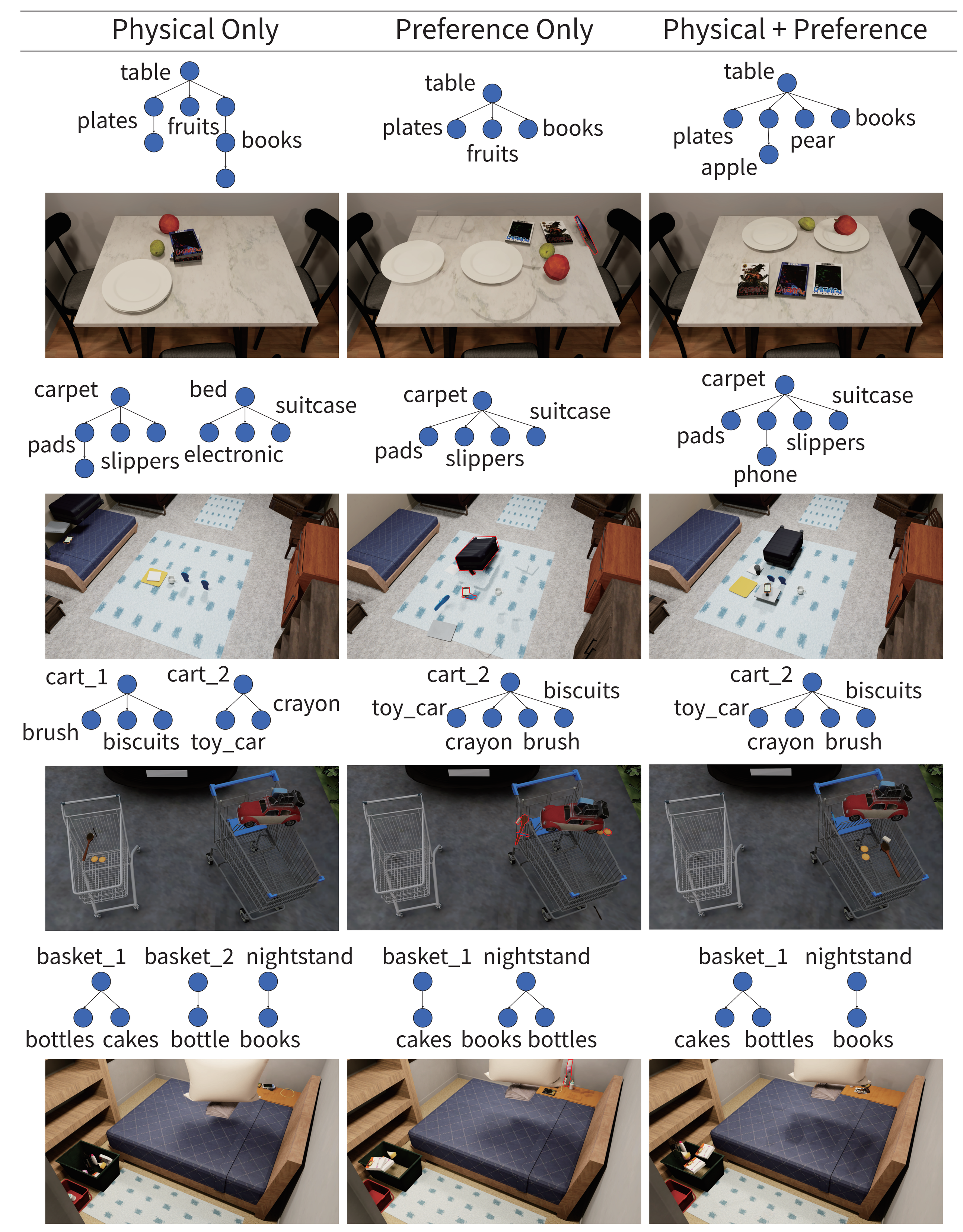
→ make objects more compact as a whole

$$\mathcal{L}_{\text{orth}} := \sigma^2(\boldsymbol{\theta})$$

→ reduce the deviation of rotation

Experiments & Results

Ablation Study. Below is a visualization of part of the results, supplemented with quantitative data analysis.

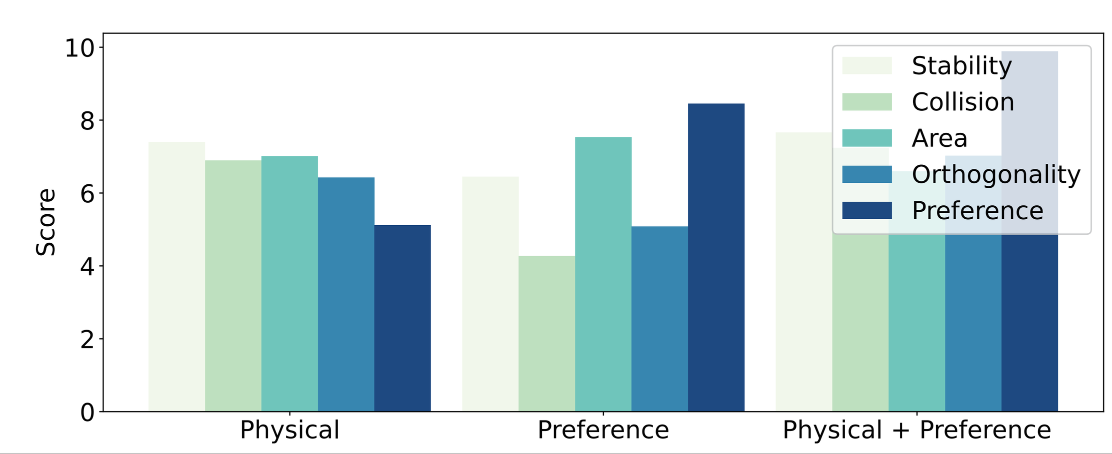


"I like that everything is laid out flat on the table, not stacked."

"The bed is for sleeping, and I don't like things that have nothing to do with sleeping on the bed."

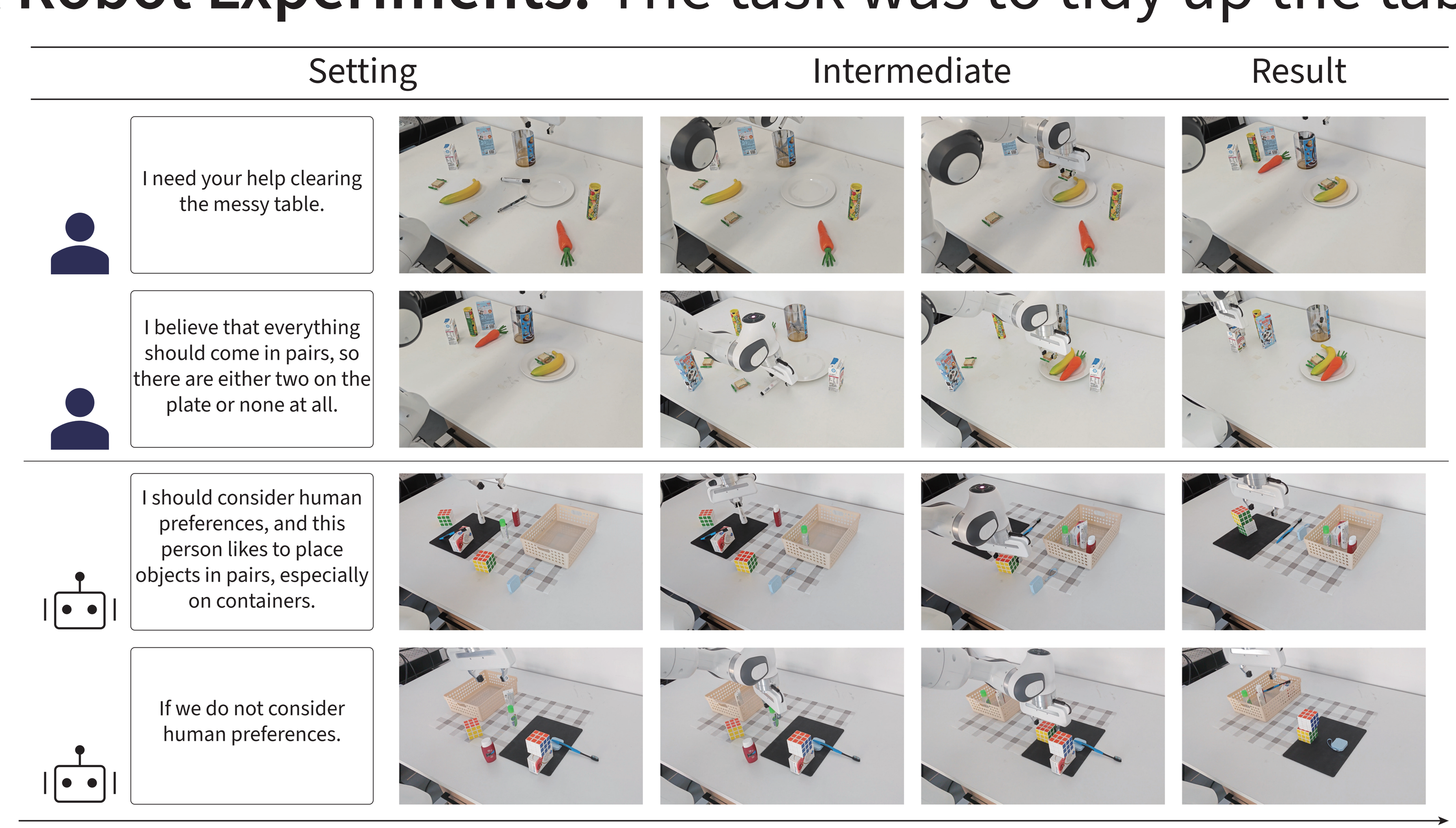
"Keep things in one place. I don't want to make two trips."

Combine the above three types of preferences.



Evaluation of preferences:
- Subjective scoring: Rated by human participants
- Objective scoring: Calculated based on the selected features

Real Robot Experiments. The task was to tidy up the table.



Conclusion

- ICLHF algorithm can learn human preferences in situ and combine them with physical constraints.
- A benchmark that incorporates human preferences into the evaluation is introduced.
- Extensive experiments were conducted to validate the effectiveness of ICLHF.

Homepage