

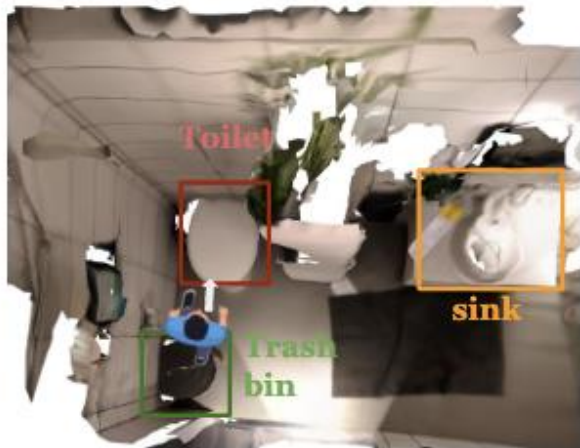


Spartun3D: Situated Spatial Understanding of 3D World in Large Language Models

Yue Zhang, Zhiyang Xu, Ying Shen,
Parisa Kordjamshidi, Lifu Huang

How to Extend to 3D World

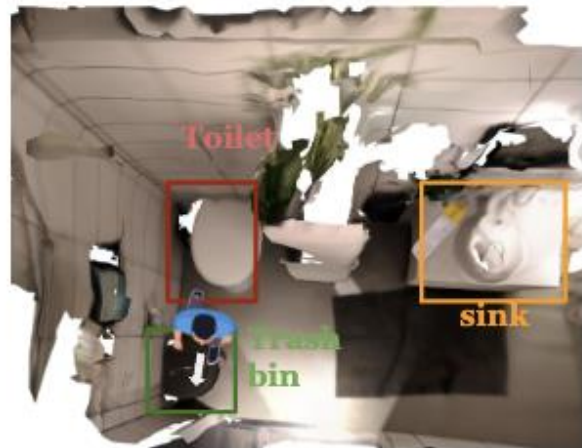
Question: what should you do to wash hands?



Situation1: You are standing beside a trash bin while there is a toilet in front of you.

3D LLMs: Scrub your hands with soap and water.

Spartun3D: You can use the sink on the right.



Situation2: You are standing beside a trash bin while there is a toilet behind you.

3D LLMs: Go to the sink and wash your hands.

Spartun3D: You can use the sink on the left.

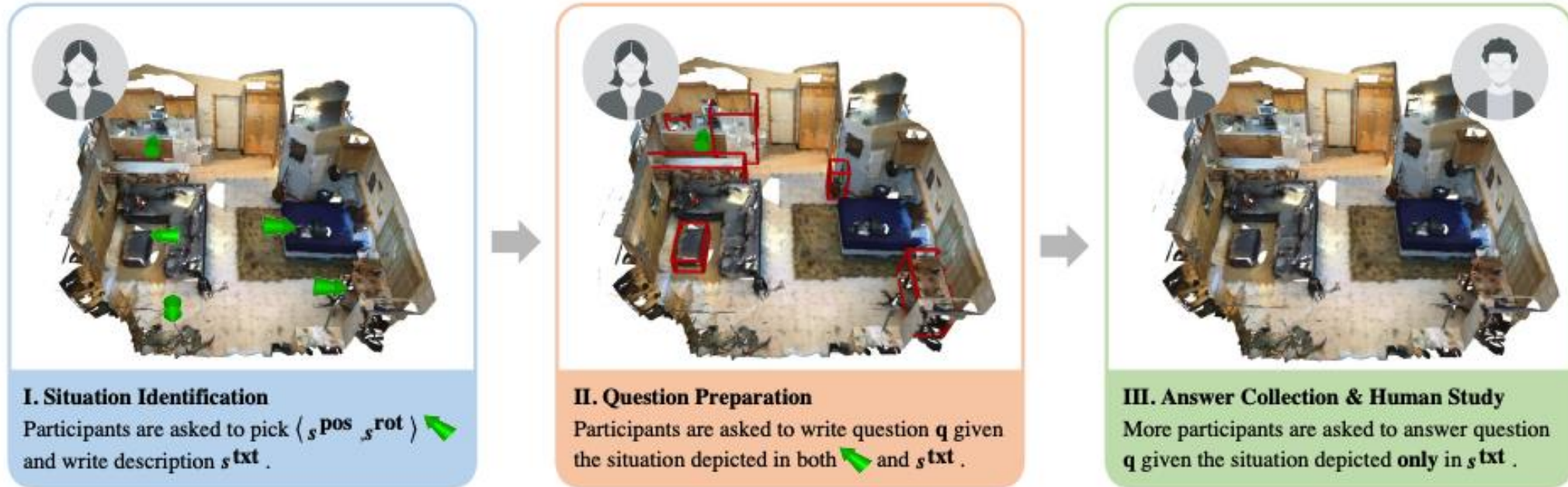
Challenge:

Current model even has trouble to understand basic situated spatial information!

Why?



Situated Question Answering in 3D Scene



Human Data Collection

Expensive

Can LLMs Help?

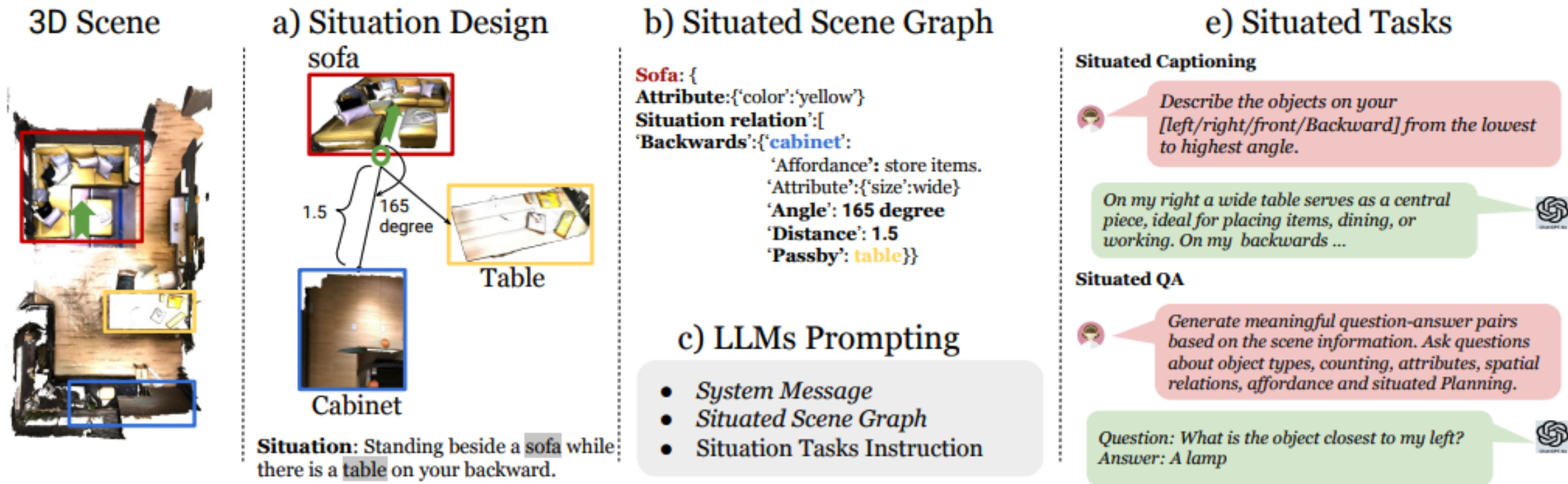


Illustration of Spartun3D Dataset Construction Process.

Spartun3D: Situated Spatial Understanding Dataset

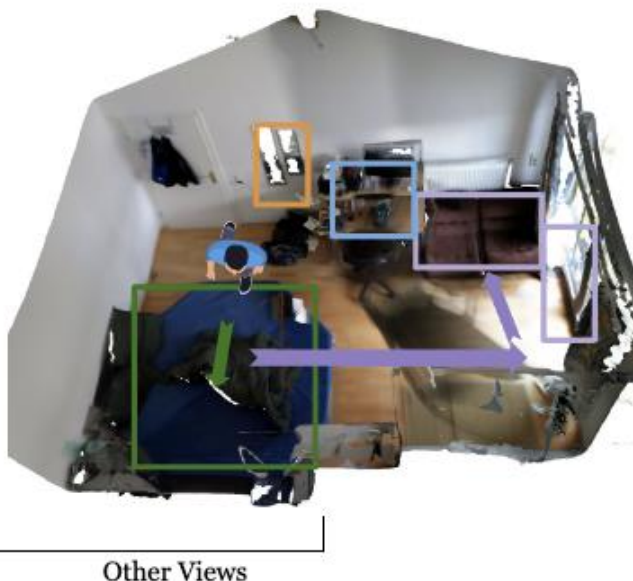
Situated Captioning

Q: Describe the scene from your current position.

A: **In front**, there's a rectangular box and a big picture that can be hung or moved. Nearby, there's a tall nightstand beside the bed. **To the right**, there's a tall pillow close to another pillow and a blanket. **Behind**, a polished mirror and a flexible bag are near a closed door with hanging clothes. **On the left**, there's an artificial lamp and a desk close to a sofa chair and trash bin, a curtain, another lamp, and a tall wardrobe near a cabinet.



Situation: Standing beside *blue octagon wide bed* that is messy while there is a *window on the left*.



Situated QA

Object Attribute and Relation

Q: Where is the desk located?

A: Behind you

Object Affordance

Q: I need to check my appearance. Where should I go?

A: You should go to the **mirror behind you**.

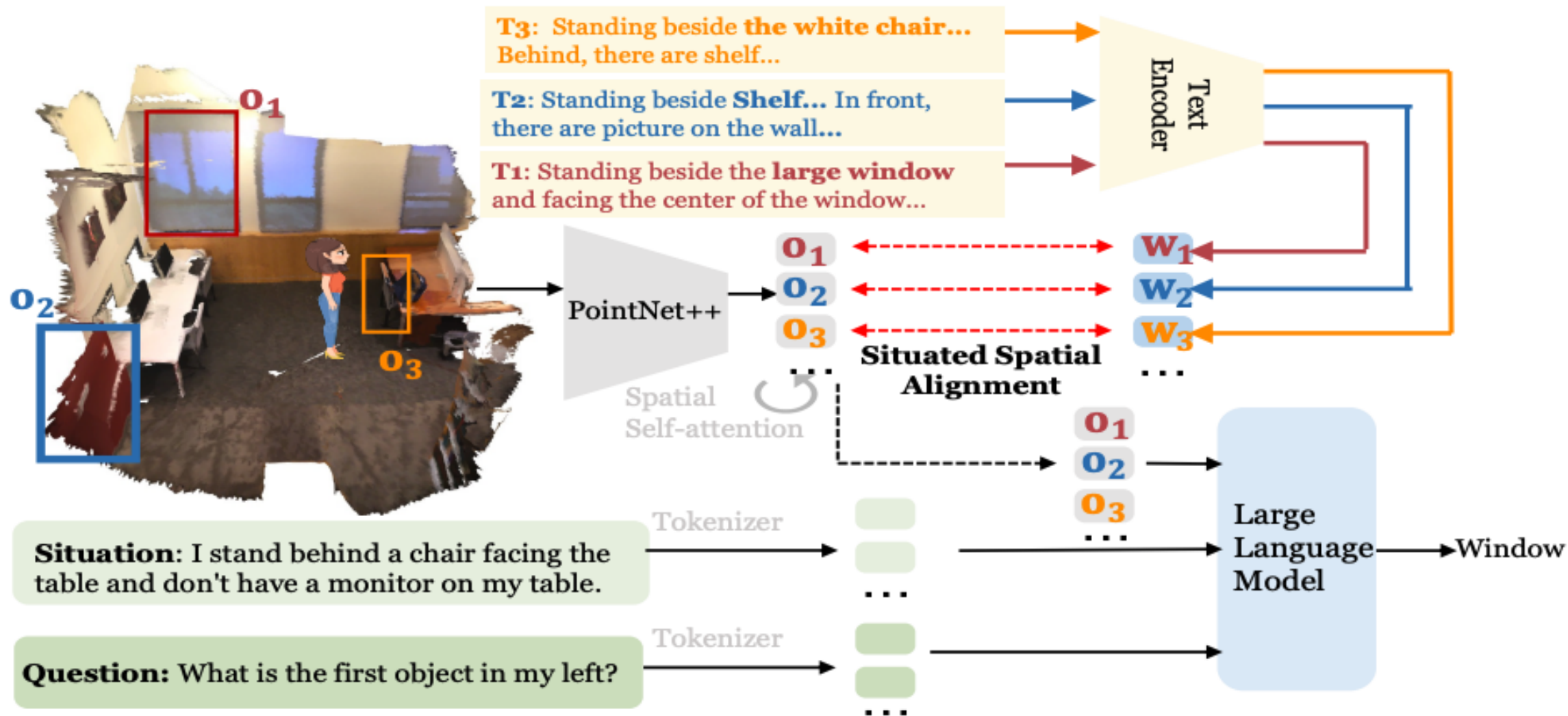
Situated Planning

Q: I want to make the room brighter and then sit down to relax. What should I do?

A: **Turn to your left** and head towards the window. Open the curtains to let in more natural light. After adjusting the curtains, head towards the **sofa chair on your left side** to sit and relax. You may need to **pass other chairs and a desk** in your path.

LLM-generated Dataset including 10K Captions and 123K QA pairs

Spartun3D-LLM



Qualitative Examples

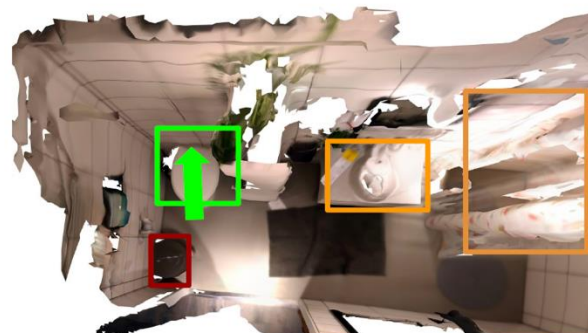


Situation: You are standing *beside sink* while there is *a potted plant* on your left.

Question: Where can you put your dirty laundry?

LEO: In the *laundry basket*.

Spa3D-LLM: In the *laundry basket on your left*.

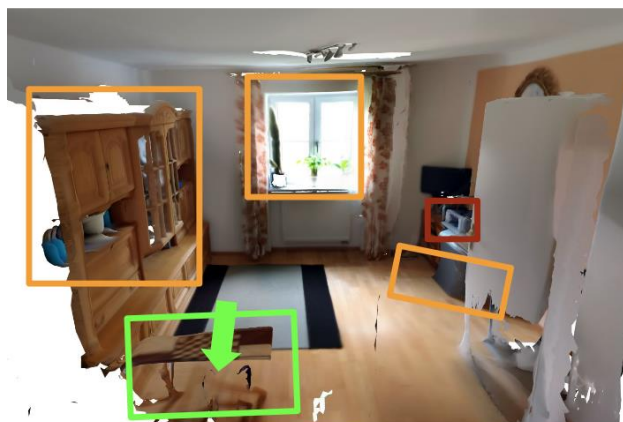


Situation: You are standing *beside toilet* while there is *a trash bin* behind you.

Question: You need to clean the sink and then adjust the shower curtain. how should you proceed ?

Spa3D-LLM (w/o align.): Turn to your right and head towards the sink to clean it. After that, *turn around* and head towards the shower curtain to adjust it.

Spa3D-LLM: Turn to your right and head towards the sink to clean it. Then, *turn slightly to your right* and head towards the shower curtain to adjust it



Situation: You are standing *beside stool* while there is *a printer* on your left.

Describe the scene from your position.

In front of me, there is no object. *To my right*, there is no object. *Behind me*, there is a wardrobe, followed by a closed rectangular glass window for looking outside. *To my left*, there's a desk for placing items in or on.

Thank You