

# *From Objects to Anywhere: A Holistic Benchmark for Multi-level Visual Grounding in 3D Scenes*



Tianxu Wang, Zhuofan Zhang, Ziyu Zhu, Yue Fan,  
Jing Xiong, Pengxiang Li, Xiaojian Ma, Qing Li

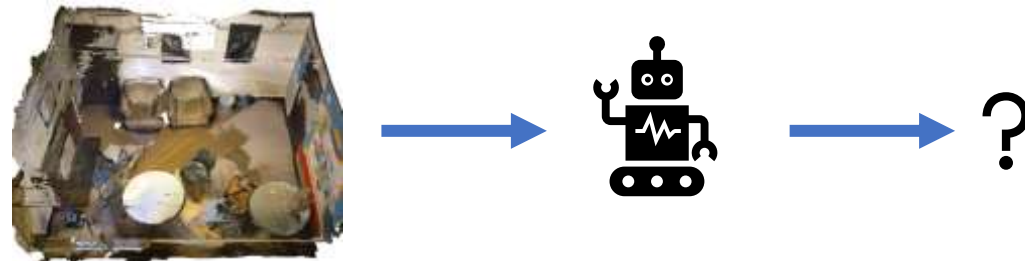


<https://anywhere-3d.github.io/>



# Motivation

- Expression: Place a floor lamp next to the armchair
- Human
  -  Visually imagine and ground the lamp in the scene
  -  Precise alignment with detailed descriptions
    - Size: 0.15 meters in base diameter, 1.2 meters in height
    - Position: adjacent to the wall and the white trash bin
  - Localize expressions to “anywhere” in 3D scenes
- Models
  - Visual grounding beyond objects?



# Anywhere3D-Bench

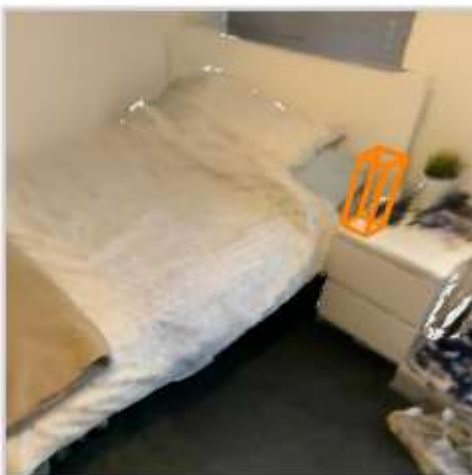
- Multi-level visual grounding in 3D scenes

**Area Level**



Choose an area that facilitates discussion and allows two people to complete the work efficiently.

**Space Level**



Lie on the bed, and place a cup with a diameter of 0.1 meters and a height of 0.3 meters at the upper right corner of the bed table.

**Object Level**



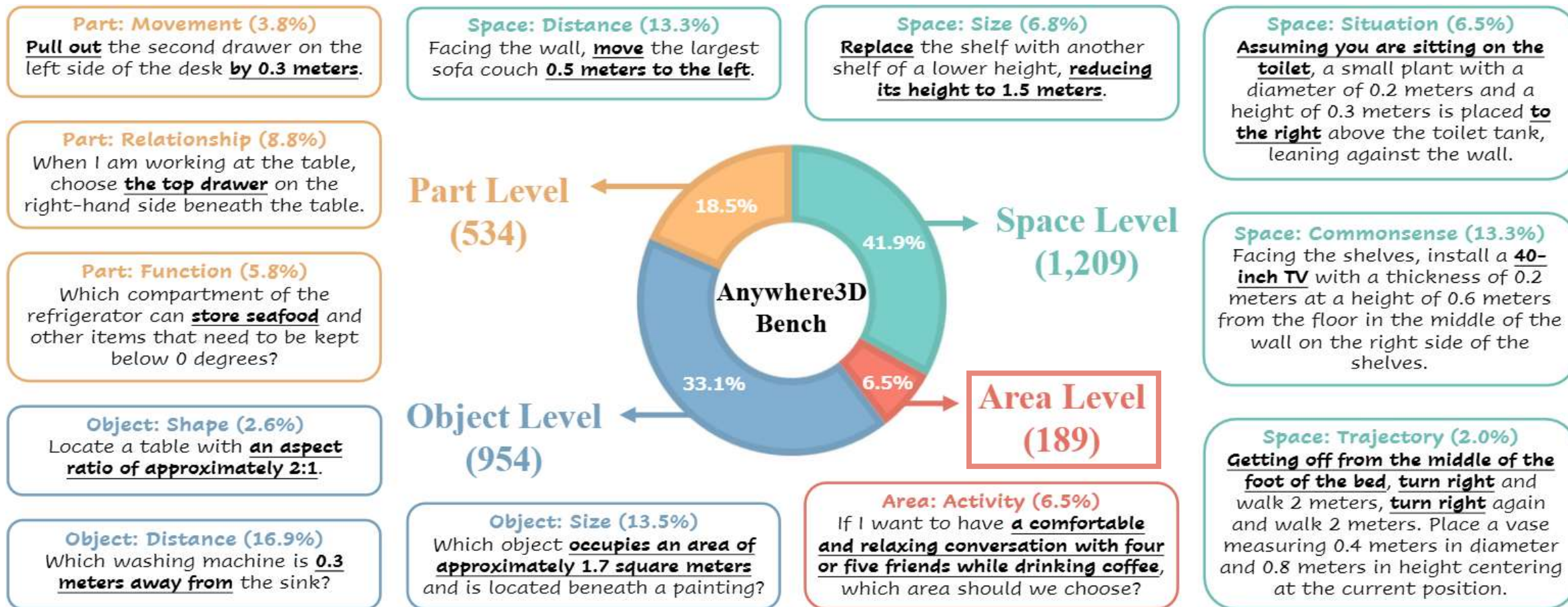
Standing up from the armchair and walking 1 meter forward, which object on the floor directly to my right is against the wall?

**Part Level**



Assume you are working at your desk. Please pull the top drawer on the left side out as far as it will go until it touches the armchair.

# Anywhere3D-Bench



# Area Level

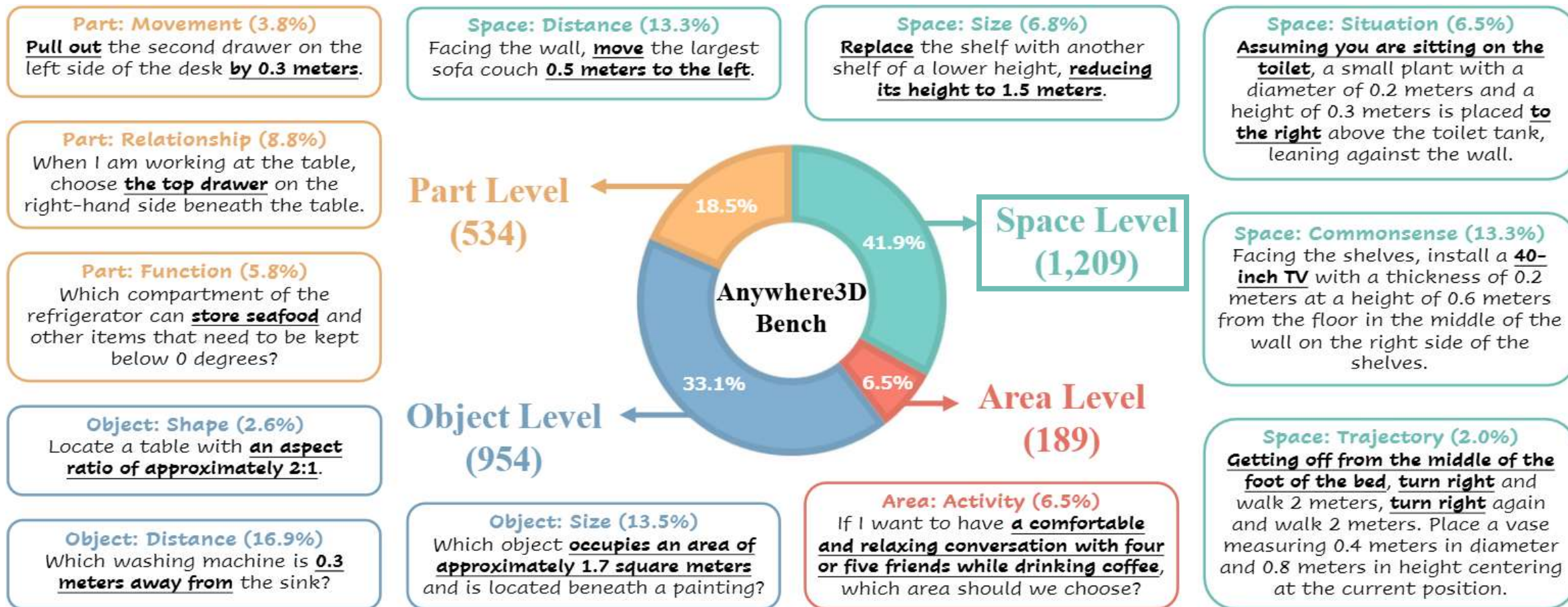


Choose the **protruding area** of the room, where you can sit on the armchair or stand in front of the curtains to gaze at the view outside the window.



Imagine you are about to deep clean the room. Identify the **area with the most cluttered object arrangement**.

# Anywhere3D-Bench



# Space Level



## Size

Replace the shelf with another shelf of a lower height, **reducing its height to 1.5 meters.**



## Distance

Face the wall, then move the largest sofa couch **0.5 meters to the left.**



## Situation

**Assuming you are sitting on the toilet**, a small plant with a diameter of 0.2 meters and a height of 0.3 meters is placed to the right above the toilet tank, leaning against the wall.

# Space Level



## Commonsense

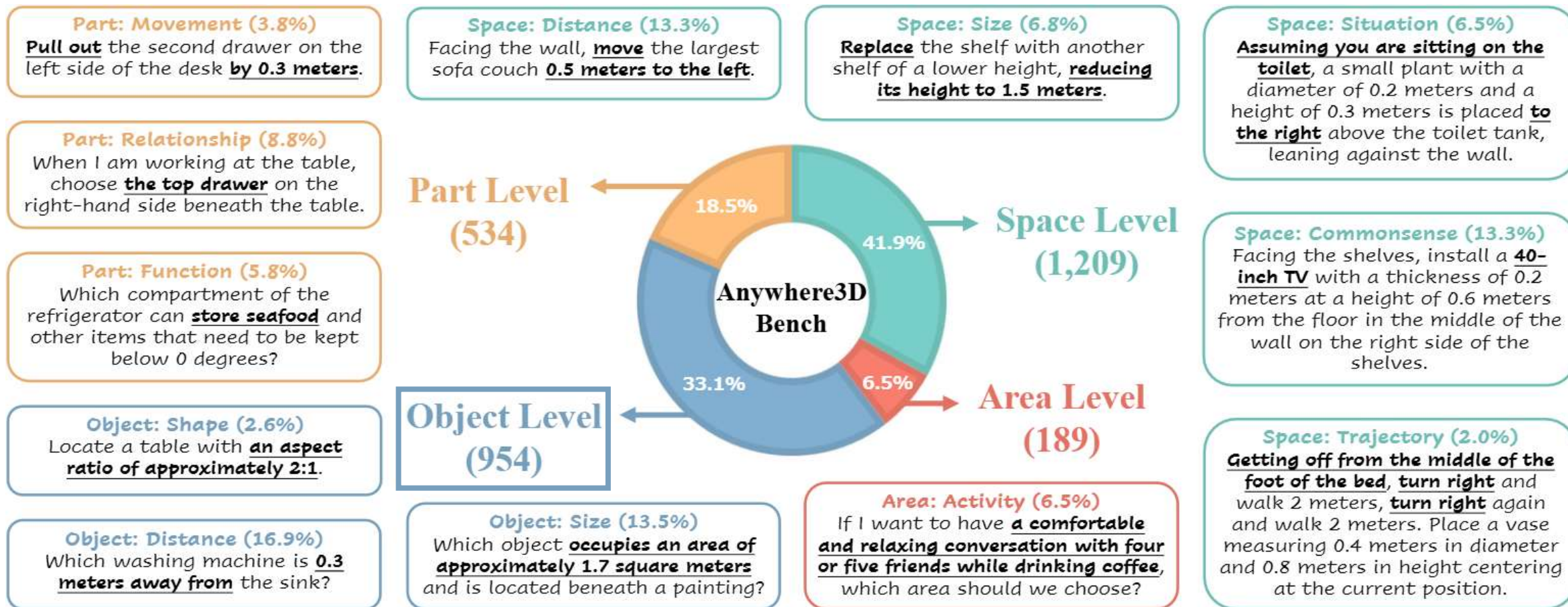
Facing the shelves, install **a 40-inch TV** with a thickness of 0.2 meters at a height of 0.6 meters from the floor in the middle of the wall on the right side of the shelves.



## Trajectory

Getting off from the middle of the foot of the bed, **turn right and walk 2 meters, turn right again and walk 2 meters**. Place a vase measuring 0.4 meters in diameter and 0.8 meters in height centering at the current position.

# Anywhere3D-Bench



# Object Level



## Shape

Locate a table **with an aspect ratio** of approximately 2:1



## Size

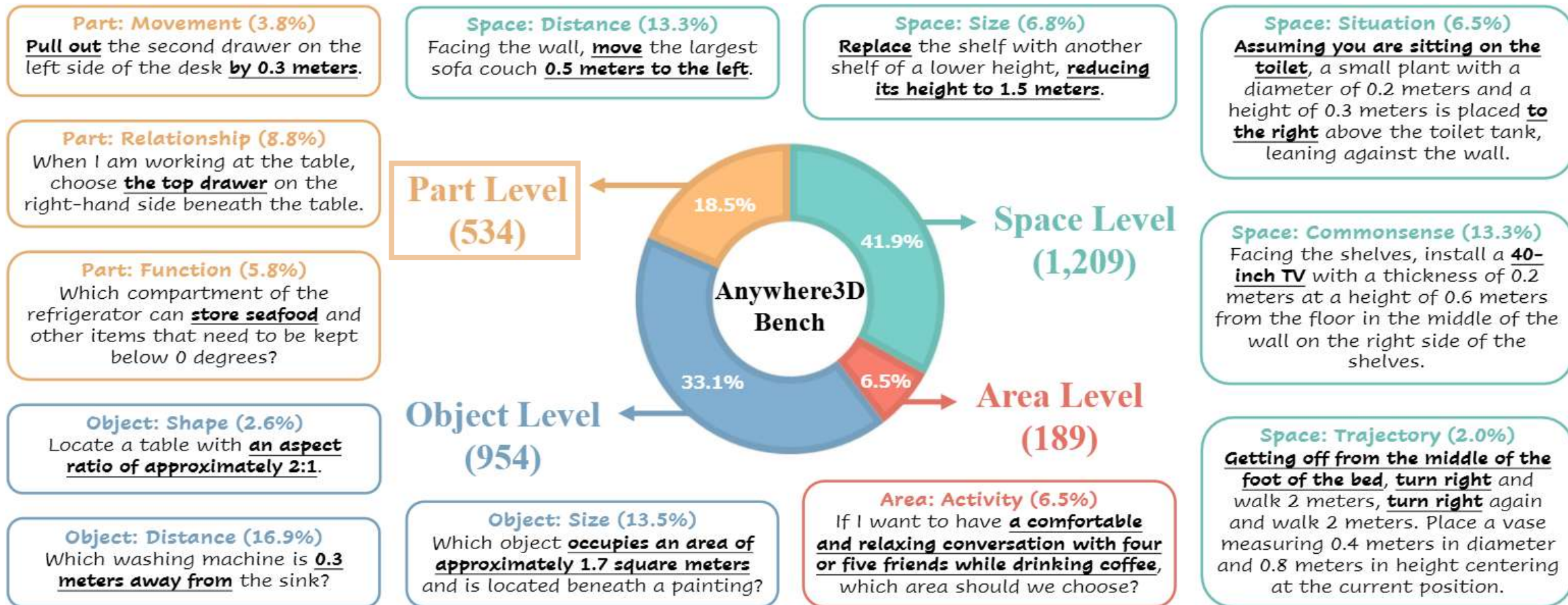
Which object occupies an area of approximately **1.7 square meters** and is located beneath a painting?



## Distance

Which washing machine **is 0.3 meters away** from the sink?

# Anywhere3D-Bench



# Part Level



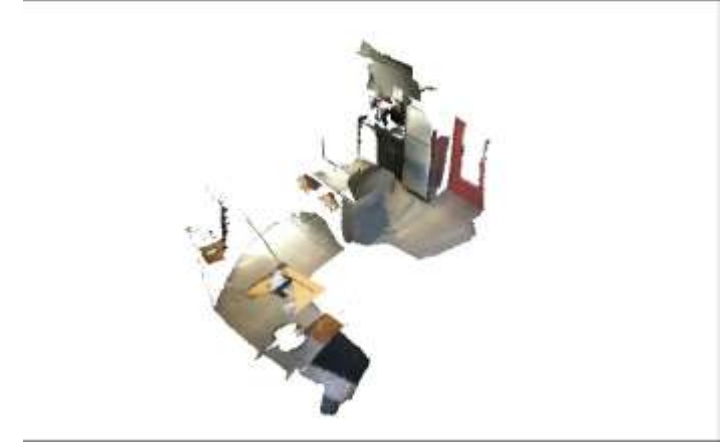
## Movement

**Pull out** the second drawer on the left side of the desk **by 0.3 meters**.



## Relationship

When I am working at the table, choose the top drawer on the **right-hand side** beneath the table.



## Function

Which compartment of the refrigerator can **store seafood** and other items that need to be kept **below 0 degrees**?

# Task Formulation

- Input  $\langle S, r \rangle$ 
  - S: 3D Scene Representation
  - r: Referring expressions for visual grounding
- Output:  $\langle \text{center\_x}, \text{center\_y}, \text{center\_z}, \text{len\_x}, \text{len\_y}, \text{len\_z} \rangle$ 
  - Predicted 3D bounding box's center coordinates and sizes along x-axis, y-axis, and z-axis

# Experimental Settings

- Benchmark LLMs, MLLMs and 3D visual grounding models

- Open-sourced models and close-sourced models
- Thinking models and non-thinking models

- Scene Representation

- LLM: textual scene graph
  - Object ground-truth locations, sizes, and captions
- MLLM:
  - 8 uniformly sampled video frames and a bird's-eye view **in addition to** LLM's textual input (following GPT4Scene's setting)
- 3D-VG models
  - Point clouds, images

```
"object_info": {  
  "object-4": {  
    "position": [1.05, -1.17, 0.68],  
    "size": [0.98, 1.24, 0.14],  
    "caption": "A possibly wooden partition with a light beige color and horizontal slat design."  
  },  
  "object-5": {  
    "position": [-0.17, -0.55, 0.07],  
    "size": [4.06, 0.43, 0.41],  
    "caption": "A smooth plaster wall section with off-white coloring and a rectangular shape. T"  
  },  
  "object-6": {  
    "position": [-1.87, 1.88, 0.58],  
    "size": [0.62, 0.54, 0.5],  
    "caption": "A possibly wooden shelf with a light beige color and horizontal slats. The struc"  
  },  
}
```

- Evaluation Metrics: Acc@0.25IoU



# Experimental Analysis: Main Results

	Open Source	Area Level	Space Level	Object Level	Part Level	Overall
<b>LLMs: <i>object bbox, captions</i></b>						
<b>non-thinking</b>						
GPT-4.1-2025-04-14	✗	<b>76.19</b>	<u>17.28</u>	<u>48.00</u>	<u>22.94</u>	<u>32.34</u>
Qwen3-32B(non-thinking)	✓	54.67	9.60	31.97	12.24	20.43
Qwen2.5-72B	✓	60.14	7.85	33.30	8.99	19.90
Qwen2.5-VL-72B	✓	56.35	6.87	29.19	9.93	18.05
DeepSeek-V3-671B-2024-12-26	✓	61.38	9.81	41.06	15.61	24.59
<b>thinking</b>						
o4-mini-2025-04-16	✗	<u>71.96</u>	<b>18.03</b>	<b>48.69</b>	<b>23.97</b>	<b>32.80</b>
Qwen3-32B(thinking)	✓	59.79	12.57	40.18	16.48	25.51
DeepSeek-R1-671B-2025-01-28	✓	<u>71.96</u>	14.61	47.76	20.92	30.49
<b>MLLMs: <i>object bbox, captions, BEV, video frames</i></b>						
<b>non-thinking</b>						
GPT-4.1-2025-04-14	✗	81.48	19.03	53.88	25.85	35.90
Gemini-2.0-Flash(non-thinking)	✗	68.43	13.18	45.39	19.97	28.70
LLaVA-NeXT-Interleave-7B	✓	6.88	0.83	4.61	2.06	2.70
LLaVA-OneVision-7B	✓	19.58	2.32	8.81	4.12	5.93
InternVL3-8B	✓	33.16	4.60	18.69	6.93	11.56
Qwen2.5-VL-72B	✓	57.16	10.56	40.74	13.80	24.19
GPT4Scene	✓	15.34	7.19	25.16	11.99	14.55
<b>thinking</b>						
o4-mini-2025-04-16	✗	76.19	23.00	55.82	31.46	38.90
o3-2025-04-16	✗	<b>87.83</b>	<b>31.26</b>	<u>60.27</u>	<b>38.77</b>	<u>45.94</u>
Gemini-2.0-Flash(thinking)	✗	81.22	21.13	53.72	28.84	37.26
Gemini-2.5-Flash	✗	81.48	23.74	54.72	30.71	39.05
Gemini-2.5-Pro	✗	<u>83.60</u>	<u>29.86</u>	<b>64.47</b>	<b>38.77</b>	<b>46.47</b>
<b>3D visual grounding models: <i>point clouds, video frames</i></b>						
PQ3D	✓	30.69	<b>8.36</b>	24.42	16.73	16.68
3D-VisTA	✓	29.10	<u>7.44</u>	25.05	15.98	16.26
Chat-Scene*	✓	<u>49.10</u>	6.55	<b>31.73</b>	<b>22.99</b>	<b>22.90</b>
Grounded 3D-LLM*	✓	<b>49.25</b>	6.62	<u>26.36</u>	<u>19.37</u>	<u>20.10</u>
Human**	–	100.00	92.00	98.00	97.00	95.00

1. Human performance substantially surpasses that of the best-performing model, **Gemini-2.5-pro** under the MLLM setting, particularly at the space level.

Space Level:  62.14%  
Overall:  48.53%

Results are presented in Acc@0.25IoU on Anywhere3D-Bench. *object bbox*: ground truth object locations and sizes. The best performance in each setting is highlighted in **bold**, and the second-best is indicated with underline.

# Experimental Analysis: Main Results

	Open Source	Area Level	Space Level	Object Level	Part Level	Overall
<b>LLMs: object bbox, captions</b>						
<b>non-thinking</b>						
GPT-4.1-2025-04-14	✗	<b>76.19</b>	<u>17.28</u>	<u>48.00</u>	<u>22.94</u>	<u>32.34</u>
Qwen3-32B(non-thinking)	✓	54.67	9.60	31.97	12.24	20.43
Qwen2.5-72B	✓	60.14	7.85	33.30	8.99	19.90
Qwen2.5-VL-72B	✓	56.35	6.87	29.19	9.93	18.05
DeepSeek-V3-671B-2024-12-26	✓	61.38	9.81	41.06	15.61	24.59
<b>thinking</b>						
o4-mini-2025-04-16	✗	<u>71.96</u>	<b>18.03</b>	<b>48.69</b>	<b>23.97</b>	<b>32.80</b>
Qwen3-32B(thinking)	✓	59.79	12.57	40.18	16.48	25.51
DeepSeek-R1-671B-2025-01-28	✓	<u>71.96</u>	14.61	47.76	20.92	30.49
<b>MLLMs: object bbox, captions, BEV, video frames</b>						
<b>non-thinking</b>						
GPT-4.1-2025-04-14	✗	81.48	19.03	53.88	25.85	35.90
Gemini-2.0-Flash(non-thinking)	✗	68.43	13.18	45.39	19.97	28.70
LLaVA-NeXT-Interleave-7B	✓	6.88	0.83	4.61	2.06	2.70
LLaVA-OneVision-7B	✓	19.58	2.32	8.81	4.12	5.93
InternVL3-8B	✓	33.16	4.60	18.69	6.93	11.56
Qwen2.5-VL-72B	✓	57.16	10.56	40.74	13.80	24.19
GPT4Scene	✓	15.34	7.19	25.16	11.99	14.55
<b>thinking</b>						
o4-mini-2025-04-16	✗	76.19	23.00	55.82	31.46	38.90
o3-2025-04-16	✗	<b>87.83</b>	<b>31.26</b>	<u>60.27</u>	<b>38.77</b>	<u>45.94</u>
Gemini-2.0-Flash(thinking)	✗	81.22	21.13	53.72	28.84	37.26
Gemini-2.5-Flash	✗	81.48	23.74	54.72	30.71	39.05
Gemini-2.5-Pro	✗	<u>83.60</u>	<u>29.86</u>	<b>64.47</b>	<b>38.77</b>	<b>46.47</b>
<b>3D visual grounding models: point clouds, video frames</b>						
PQ3D	✓	30.69	<b>8.36</b>	24.42	16.73	16.68
3D-VisTA	✓	29.10	<u>7.44</u>	25.05	15.98	16.26
Chat-Scene*	✓	<u>49.10</u>	6.55	<b>31.73</b>	<b>22.99</b>	<b>22.90</b>
Grounded 3D-LLM*	✓	<b>49.25</b>	6.62	<u>26.36</u>	<u>19.37</u>	<u>20.10</u>
Human**	–	100.00	92.00	98.00	97.00	95.00

2. Area v.s. Space v.s. Object v.s. Part

**Challenging space-level task:** understand spatial relations, distances, situations in 3D scenes beyond objects

**Challenging part-level task:** identify the object first, then reason about the part's location and size based on relationships, functions.

Results are presented in Acc@0.25IoU on Anywhere3D-Bench. *object bbox*: ground truth object locations and sizes. The best performance in each setting is highlighted in **bold**, and the second-best is indicated with underline.

# Experimental Analysis: Main Results

	Open Source	Area Level	Space Level	Object Level	Part Level	Overall
<b>LLMs: <i>object bbox, captions</i></b>						
<b>non-thinking</b>						
GPT-4.1-2025-04-14	✗	<b>76.19</b>	<u>17.28</u>	<u>48.00</u>	<u>22.94</u>	<u>32.34</u>
Qwen3-32B(non-thinking)	✓	54.67	9.60	31.97	12.24	20.43
Qwen2.5-72B	✓	60.14	7.85	33.30	8.99	19.90
Qwen2.5-VL-72B	✓	56.35	6.87	29.19	9.93	18.05
DeepSeek-V3-671B-2024-12-26	✓	61.38	9.81	41.06	15.61	24.59
<b>thinking</b>						
o4-mini-2025-04-16	✗	<u>71.96</u>	<b>18.03</b>	<b>48.69</b>	<b>23.97</b>	<b>32.80</b>
Qwen3-32B(thinking)	✓	59.79	12.57	40.18	16.48	25.51
DeepSeek-R1-671B-2025-01-28	✓	<u>71.96</u>	14.61	47.76	20.92	30.49
<b>MLLMs: <i>object bbox, captions, BEV, video frames</i></b>						
<b>non-thinking</b>						
GPT-4.1-2025-04-14	✗	81.48	19.03	53.88	25.85	35.90
Gemini-2.0-Flash(non-thinking)	✗	68.43	13.18	45.39	19.97	28.70
LLaVA-NeXT-Interleave-7B	✓	6.88	0.83	4.61	2.06	2.70
LLaVA-OneVision-7B	✓	19.58	2.32	8.81	4.12	5.93
InternVL3-8B	✓	33.16	4.60	18.69	6.93	11.56
Qwen2.5-VL-72B	✓	57.16	10.56	40.74	13.80	24.19
GPT4Scene	✓	15.34	7.19	25.16	11.99	14.55
<b>thinking</b>						
o4-mini-2025-04-16	✗	76.19	23.00	55.82	31.46	38.90
o3-2025-04-16	✗	<b>87.83</b>	<b>31.26</b>	<u>60.27</u>	<b>38.77</b>	<u>45.94</u>
Gemini-2.0-Flash(thinking)	✗	81.22	21.13	53.72	28.84	37.26
Gemini-2.5-Flash	✗	81.48	23.74	54.72	30.71	39.05
Gemini-2.5-Pro	✗	<u>83.60</u>	<u>29.86</u>	<b>64.47</b>	<b>38.77</b>	<b>46.47</b>
<b>3D visual grounding models: <i>point clouds, video frames</i></b>						
PQ3D	✓	30.69	<b>8.36</b>	24.42	16.73	16.68
3D-VisTA	✓	29.10	7.44	25.05	15.98	16.26
Chat-Scene*	✓	<u>49.10</u>	6.55	<b>31.73</b>	<b>22.99</b>	<b>22.90</b>
Grounded 3D-LLM*	✓	<b>49.25</b>	6.62	<u>26.36</u>	<u>19.37</u>	<u>20.10</u>
Human**	–	100.00	92.00	98.00	97.00	95.00

## 3. MLLMs v.s. 3D models

Specialized 3D visual grounding models demonstrate limited performance, particularly **on space level**. Current 3D-VG models lack generalizability for **multi-level** visual grounding tasks

Overall: ↗ 23.57%

Results are presented in Acc@0.25IoU on Anywhere3D-Bench. *object bbox*: ground truth object locations and sizes.

The best performance in each setting is highlighted in **bold**, and the second-best is indicated with underline.

# Experimental Analysis: Main Results

	Open Source	Area Level	Space Level	Object Level	Part Level	Overall
<b>LLMs: object bbox, captions</b>						
<b>non-thinking</b>						
GPT-4.1-2025-04-14	✗	<b>76.19</b>	<u>17.28</u>	<u>48.00</u>	<u>22.94</u>	<u>32.34</u>
Qwen3-32B(non-thinking)	✓	54.67	9.60	31.97	12.24	20.43
Qwen2.5-72B	✓	60.14	7.85	33.30	8.99	19.90
Qwen2.5-VL-72B	✓	56.35	6.87	29.19	9.93	18.05
DeepSeek-V3-671B-2024-12-26	✓	61.38	9.81	41.06	15.61	24.59
<b>thinking</b>						
o4-mini-2025-04-16	✗	71.96	<b>18.03</b>	<b>48.69</b>	<b>23.97</b>	<b>32.80</b>
Qwen3-32B(thinking)	✓	59.79	<u>12.57</u>	<u>40.18</u>	<u>16.48</u>	<u>25.51</u>
DeepSeek-R1-671B-2025-01-28	✓	<u>71.96</u>	14.61	47.76	20.92	30.49
<b>MLLMs: object bbox, captions, BEV, video frames</b>						
<b>non-thinking</b>						
GPT-4.1-2025-04-14	✗	81.48	19.03	53.88	25.85	35.90
Gemini-2.0-Flash(non-thinking)	✗	68.43	13.18	45.39	19.97	28.70
LLaVA-NeXT-Interleave-7B	✓	6.88	0.83	4.61	2.06	2.70
LLaVA-OneVision-7B	✓	19.58	2.32	8.81	4.12	5.93
InternVL3-8B	✓	33.16	4.60	18.69	6.93	11.56
Qwen2.5-VL-72B	✓	57.16	10.56	40.74	13.80	24.19
GPT4Scene	✓	15.34	7.19	25.16	11.99	14.55
<b>thinking</b>						
o4-mini-2025-04-16	✗	76.19	23.00	55.82	31.46	38.90
o3-2025-04-16	✗	<b>87.83</b>	<b>31.26</b>	<u>60.27</u>	<b>38.77</b>	45.94
Gemini-2.0-Flash(thinking)	✗	81.22	21.13	53.72	28.84	37.26
Gemini-2.5-Flash	✗	81.48	23.74	54.72	30.71	39.05
Gemini-2.5-Pro	✗	<u>83.60</u>	<u>29.86</u>	<b>64.47</b>	<b>38.77</b>	<b>46.47</b>
<b>3D visual grounding models: point clouds, video frames</b>						
PQ3D	✓	30.69	<b>8.36</b>	24.42	16.73	16.68
3D-VisTA	✓	29.10	<u>7.44</u>	25.05	15.98	16.26
Chat-Scene*	✓	<u>49.10</u>	6.55	<b>31.73</b>	<b>22.99</b>	<b>22.90</b>
Grounded 3D-LLM*	✓	<b>49.25</b>	6.62	<u>26.36</u>	<u>19.37</u>	<u>20.10</u>
Human**	–	100.00	92.00	98.00	97.00	95.00

Qwen3-32B: ↑ 5.08%

4. Thinking v.s. Non-thinking  
Reasoning capabilities are required to effectively addressing Anywhere3D-bench

Gemini-2.0-Flash: ↑ 8.56%

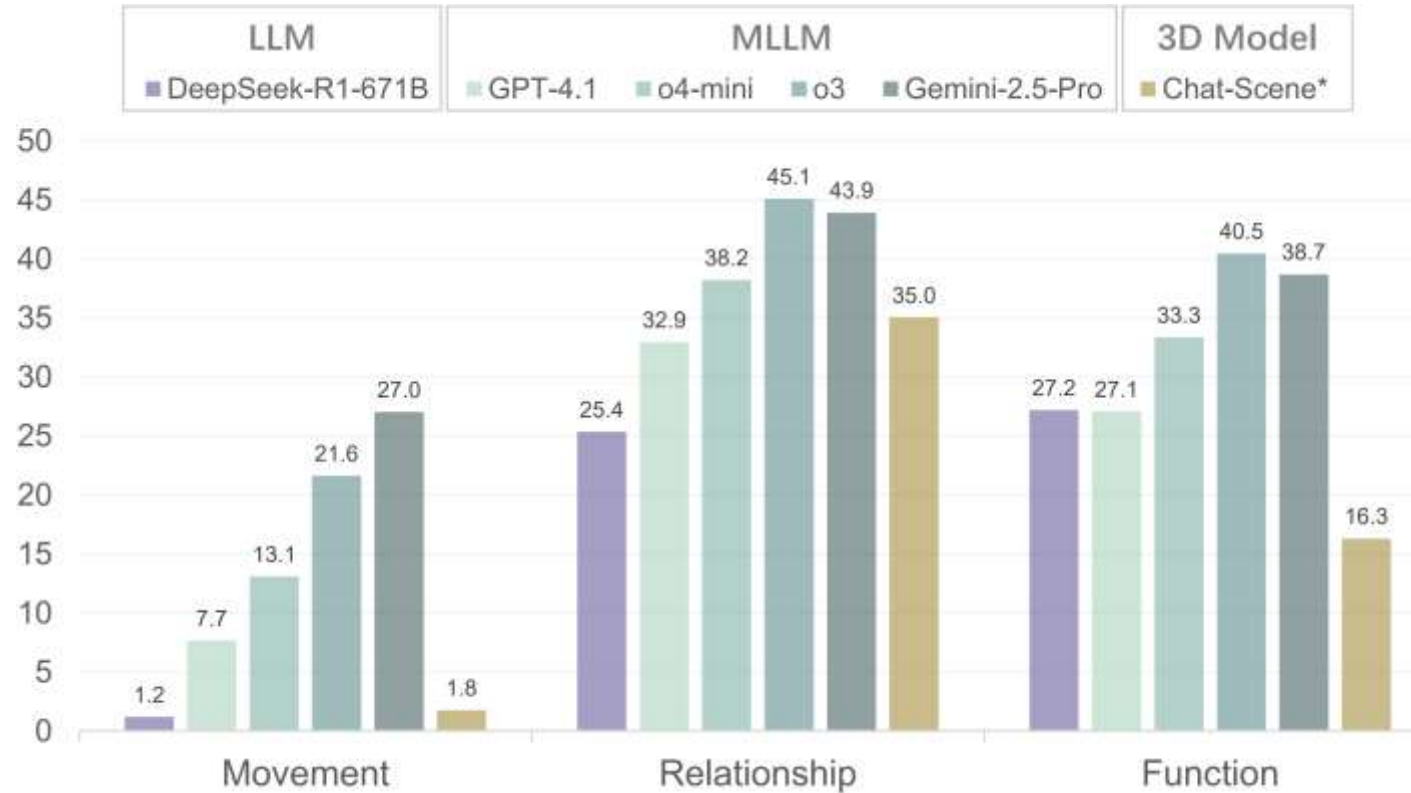
Results are presented in Acc@0.25IoU on Anywhere3D-Bench. *object bbox*: ground truth object locations and sizes. The best performance in each setting is highlighted in **bold**, and the second-best is indicated with underline.

# Experimental Analysis: Space Level Detailed Analysis



- Trajectory: Most challenging task at space-level  
Comprehensive understanding of spatial distance, relationships, and orientation
- Size: most straightforward  
Easier for models to adjust objects' sizes and minor positional refinements

# Experimental Analysis: Part Level Detailed Analysis



- Dynamic Movement: most challenging task at part-level
  - Visual ability: Identify the **correct part** of the object
  - Spatial ability: Then predict the bounding box **after the movement** by understanding object orientation / spatial relationship

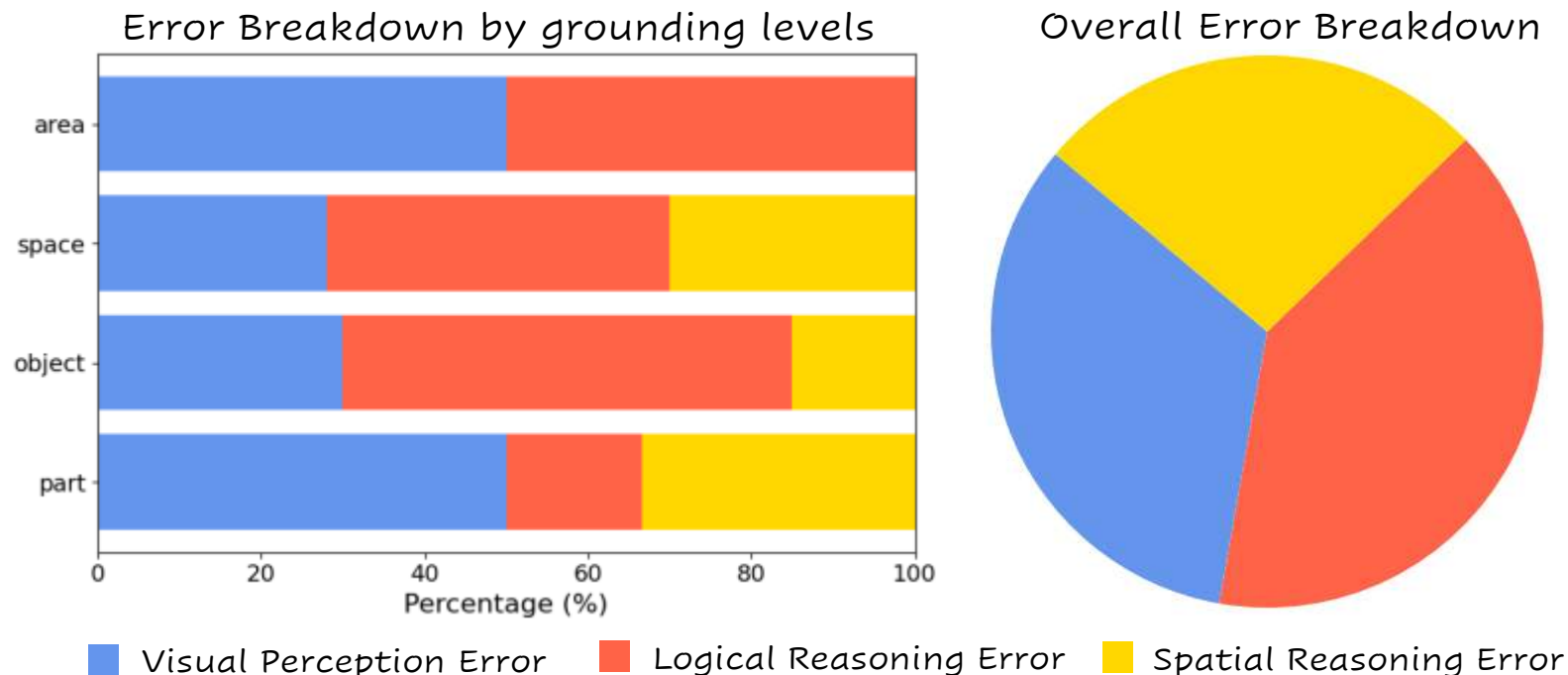
# Error Breakdown

- Error analysis of the best-performing model Gemini-2.5-pro on human-evaluation subset

**Visual Perception Error:** Arising from the misrecognition of object categories, attributes, or other visual properties based on visual inputs or object captions.

**Logical Reasoning Error:** Referring to errors stemming from incorrect logical reasoning, including incomplete comprehension of the referring expression or commonsense reasoning failures not attributable to visual perception.

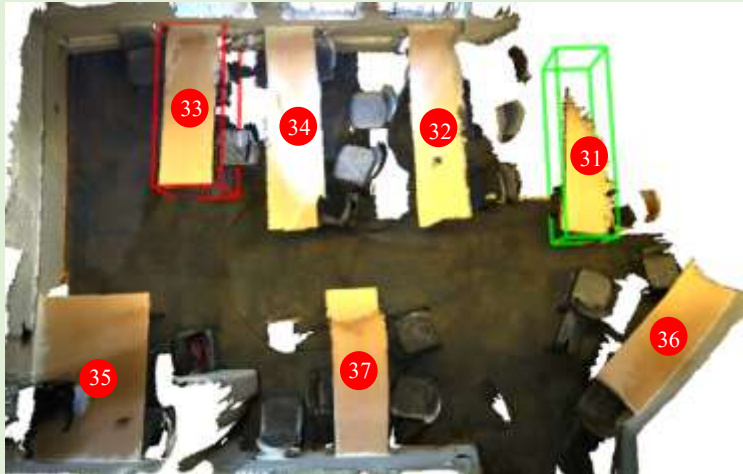
**Spatial Reasoning Error:** Arising from incorrect identification of object orientation or inaccurate alignment between egocentric descriptions and the global coordinate system (x, y, z).



# Qualitative Results

## Referring Expression

Find the table approximately 0.9 meters away from the slanted-positioned table.



### Gemini-2.5-Pro Reasoning Process

1. Identify the 'slanted-positioned table': analyze the Bird's Eye View (BEV) image to locate a table that is positioned at an angle relative to the other tables and the room's layout. The tables 'object-31', 'object-32', '**object-36**', and 'object-37' are all **aligned with the primary axes of the room**. In contrast, '**object-35**' is clearly positioned at a **diagonal, making it the "slanted-positioned table"**



**Visual Perception Error**

(a)

## Referring Expression

Standing up from the middle of the bed and walk 1.1 meters forward. Turn left and walk another 2.8 meters. Place a floor lamp measuring 0.4 meters in base diameter and 1.2 meters in height centered at your current position.



### Gemini-2.5-Pro Reasoning Process

1. **\*\*Deconstruct the referring expression:** Starting Point: "Standing up from the middle of the bed". The bed is identified as 'object-3'. From the scene graph, the position of the bed's center is '[0.67, 0.77, 0.49]'. **We will use the horizontal coordinates '(x, y) = (0.67, 0.77)' as the starting point.**

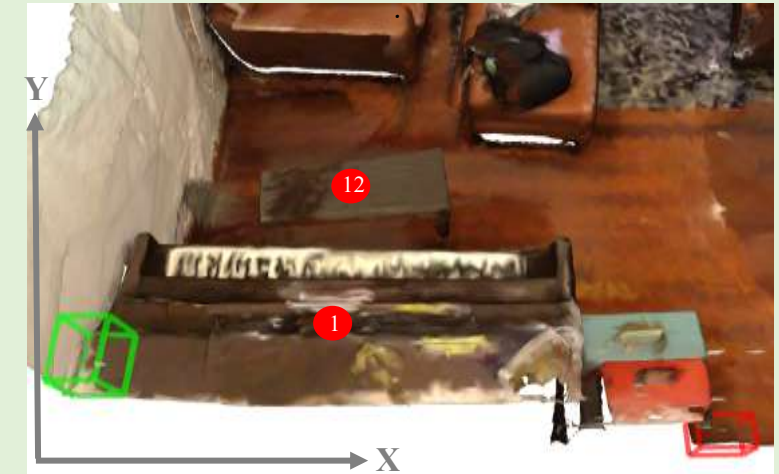


**Logical Reasoning Error**

(b)

## Referring Expression

Facing the piano, place a small speaker measuring 0.2 meters in length, 0.13 meters in width, and 0.2 meters in height immediately next to the wall on the far right of the piano, ensuring easy listening while seated on the piano bench.



### Gemini-2.5-Pro Reasoning Process

1. Reference Frame: The expression "Facing the piano" establishes the orientation. The piano (object-1) is positioned with its keyboard facing the positive y-direction. A person sitting on the piano bench (object-12) would be facing the piano, i.e., looking in the negative y-direction. From this perspective, **"right" corresponds to the positive x-direction.**



**Spatial Reasoning Error**

(c)

Fig: Qualitative Results. **Green bounding boxes** represent ground-truth while **red boxes** represent Gemini-2.5-pro's prediction. The error in reasoning process made by Gemini-2.5-pro is highlighted in **bold**.

# Non-thinking model v.s. thinking model

- Comparison between the best-performing non-thinking model(GPT-4.1) and thinking model(Gemini-2.5-pro)

## Referring Expression

Imagine you are facing the door, and install an antique clock with a diameter of 0.4 meters and a thickness of 0.15 meters directly above the shelves on the left side of the gate. The clock is positioned 0.3 meters above the center of the top of the shelves, complementing the metallic texture of the shelves perfectly.

GPT-4.1-2025-04-14



**\*\*Step 1: Identify shelves' position and size:\*\***  
**\*\*Step 2: Positioning the clock:\*\***  
The clock should be "0.3 meters above the center of the top of the shelves":

- $x = -1.21$
- $y = 1.59$
- $z = 2.07 + 0.3 = 2.37$
- **\*\*Step 3: Clock size:\*\***
  - **Diameter = 0.4, so  $xlength = ylength = 0.4$  Thickness = 0.15, so  $zlength = 0.15$**



Gemini-2.5-Pro



**\*\*Step 1: Calculate the clock's center point:**  
positioned "0.3 meters above" --- This means adding 0.3 meters to the z-coordinate.

**\*\*Step 2: Determine the Clock's size:\*\***  
The shelves are oriented along a wall parallel to the YZ plane (indicated by its small x-dimension). **A wall-mounted clock's "thickness" would be its dimension perpendicular to the wall, which is along the x-axis.** The "diameter" applies to the face of the clock, covering the y and z dimensions.  
So  $xlength(thickness) = 0.15$ ,  $ylength = zlength = 0.4$



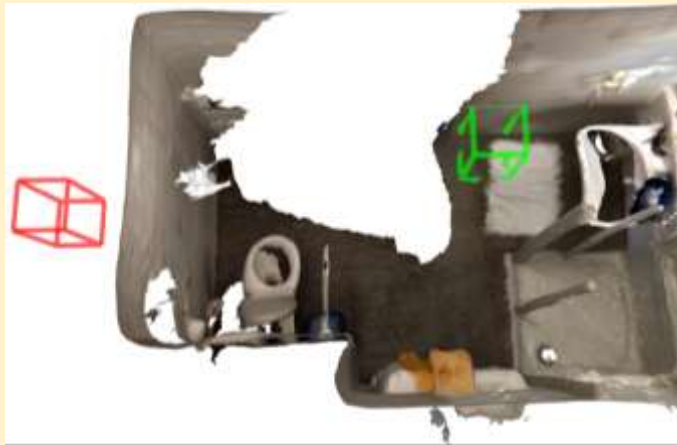
# Non-thinking model v.s. thinking model

- Comparison between the best-performing non-thinking model(GPT-4.1) and thinking model(Gemini-2.5-pro)

## Referring Expression

Getting up from the toilet and facing the opposite wall, walk 0.6 meters ahead and turn right. Then walk 1.6 meters ahead and stop. Position a resting chair, with base diameter 0.35 meters and height 0.8 meters, centered at the current position.

GPT-4.1-2025-04-14



**\*\*Step 1: Starting Point\*\*** --- The toilet is object-9. **Its center is at [1.45, 0.03, 0.79]**

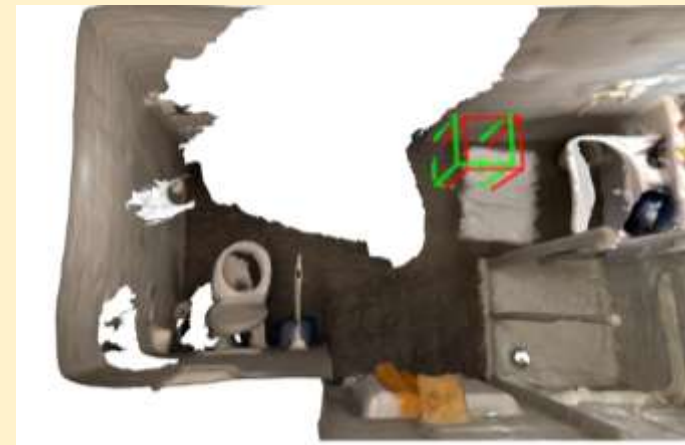
**\*\*Step 2: Facing the opposite wall (from the toilet)\*\*:** - The toilet is set along the right wall (positive x, near y=0). - "Opposite wall" means facing roughly the negative y-direction.

**\*\*Step 3: Walk 0.6 meters ahead (negative y direction):\*\*** - Advance from (1.45, 0.03) to (1.45, 0.03 - 0.6) = (1.45, -0.57)

**\*\*Step 4: Turn Right:** **now facing positive x direction, the forward direction is along +x.**



Gemini-2.5-Pro



**\*\*Step 1: Identify the starting object and position\*\*** --- From the BEV image, the toilet is positioned against the wall and faces towards the negative y-axis. **"Getting up" implies the starting point is assumed to be at the front edge of the toilet's bounding box.** The y-coordinate of the front edge is  $\text{toilet\_y\_center} - \text{toilet\_y\_size} / 2 = 0.03 - 0.74 / 2 = -0.34$ .

**\*\*Step 2: Determine initial orientation:\*\*** along the negative y-axis.

**\*\*Step 3: Simulate Movement:\*\*** **Turning right from facing the negative y-axis means the new direction is along the negative x-axis**



Please scan the QR code or visit our website for more information.

<https://anywhere-3d.github.io/>

