

# FunFact: Building Probabilistic Functional 3D Scene Graphs via Factor-Graph Reasoning

Zhengyu Fu<sup>1</sup>, René Zurbrügg<sup>1</sup>, Kaixian Qu<sup>1</sup>, Marc Pollefeys<sup>1,2</sup>, Marco Hutter<sup>1</sup>, Hermann Blum<sup>3</sup>, Zuria Bauer<sup>1</sup>

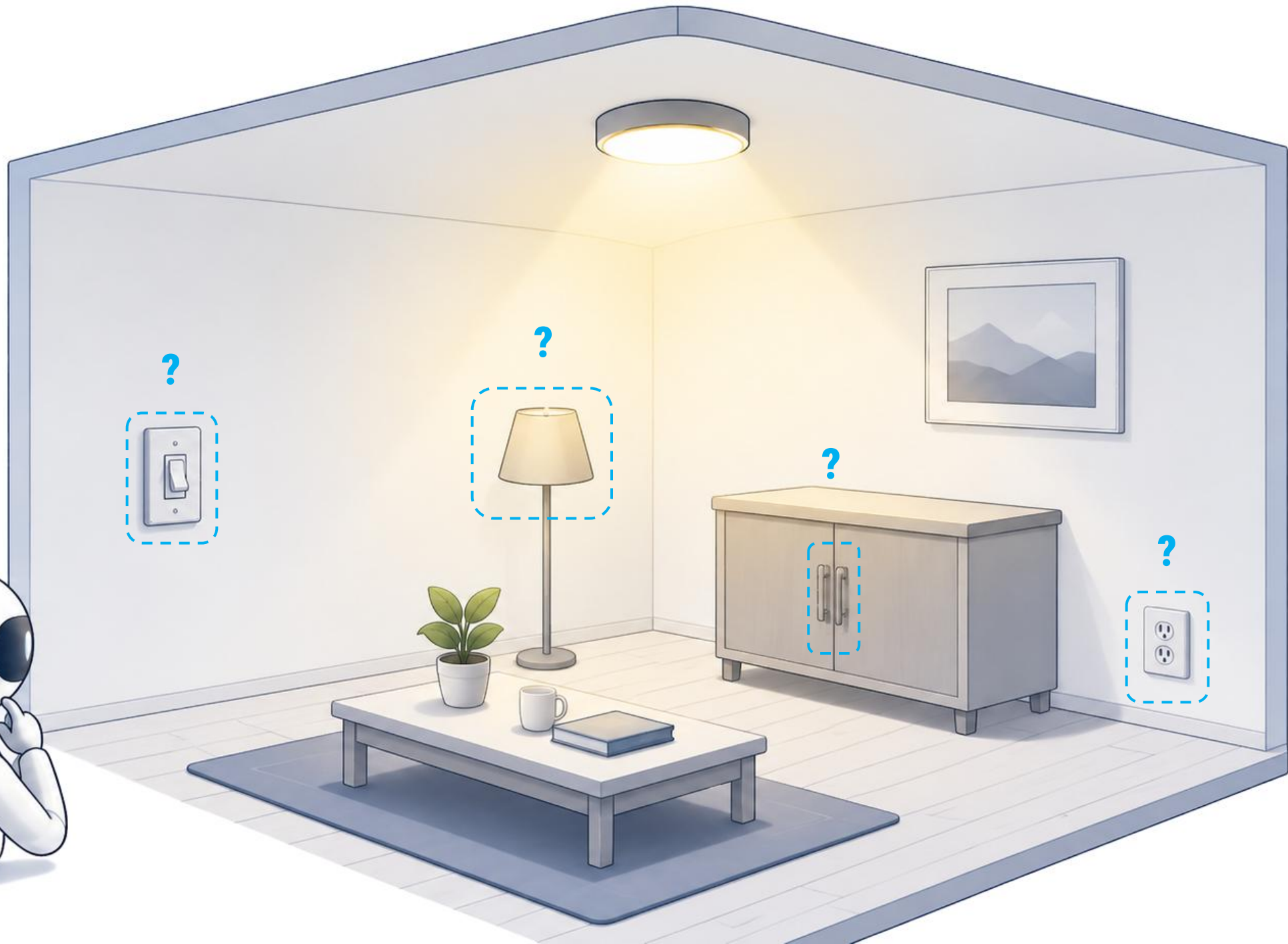
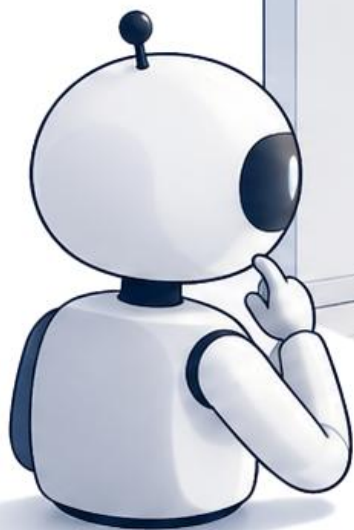
<sup>1</sup>ETH Zürich

<sup>2</sup>Microsoft

<sup>3</sup>University of Bonn & Lamar Institute

# Motivation

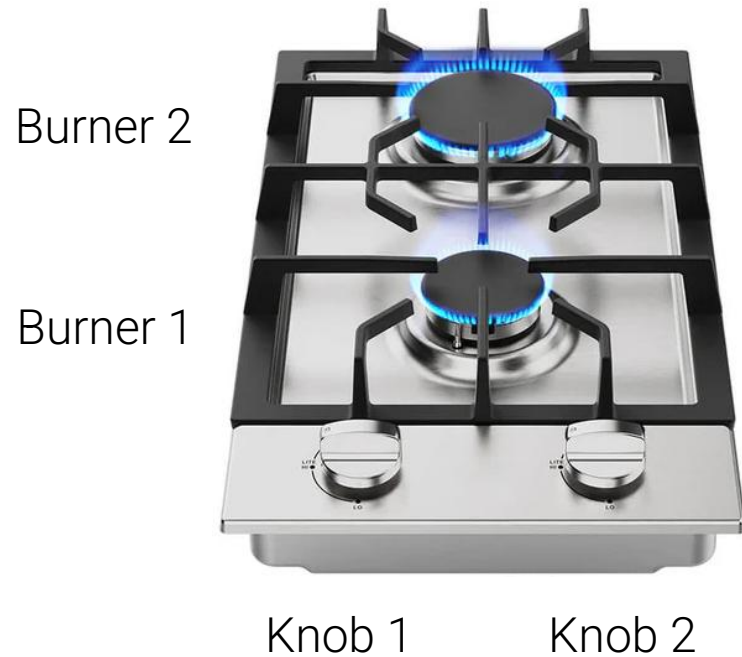
Which object should I interact with to change the illumination?



# Challenges

1. Many functional connections are **visually ambiguous**.

Which knob controls which burner ?



Knob 1 --> Burner 1  
Knob 2 --> Burner 2

**Or**

Knob 1 --> Burner 2  
Knob 2 --> Burner 1

# Challenges

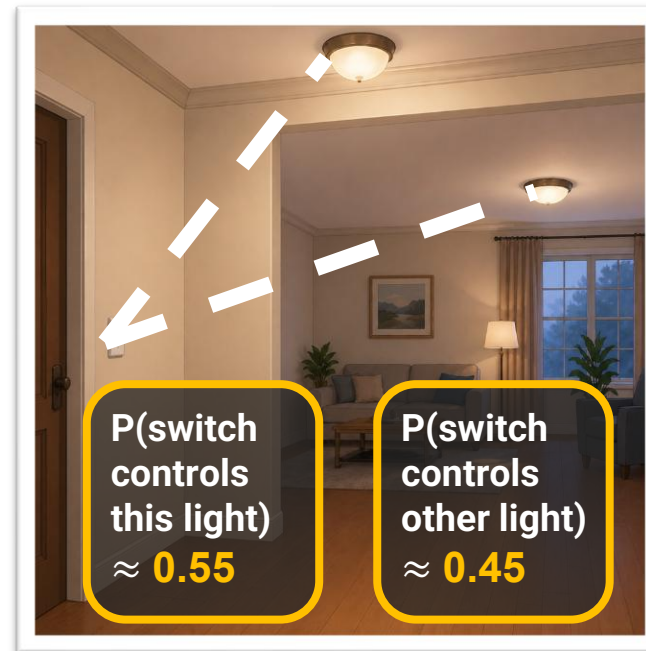
## 2. Functional connections are **interdependent**.

Scene-wide context changes per-edge confidence estimates.

**Entering an unknown room**  
switch + one ceiling light



**Turning around**  
Another ceiling light



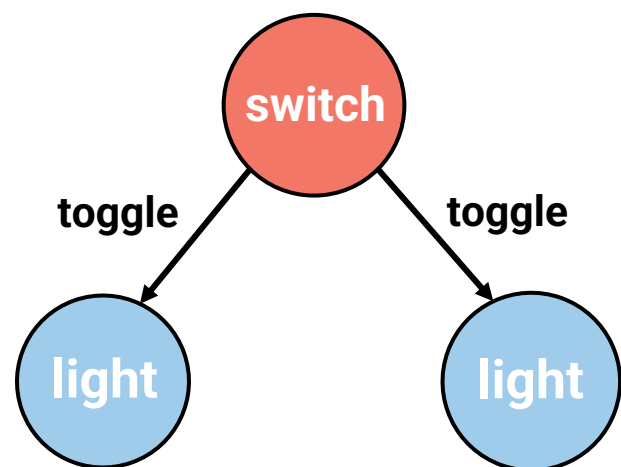
High confidence  
for a single connection

discovering  
more context

Confidence drops  
for both connections

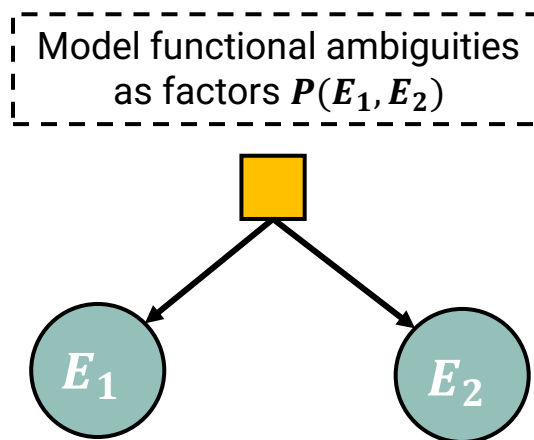
# Core Idea

## Functional Scene Graph (FSG)



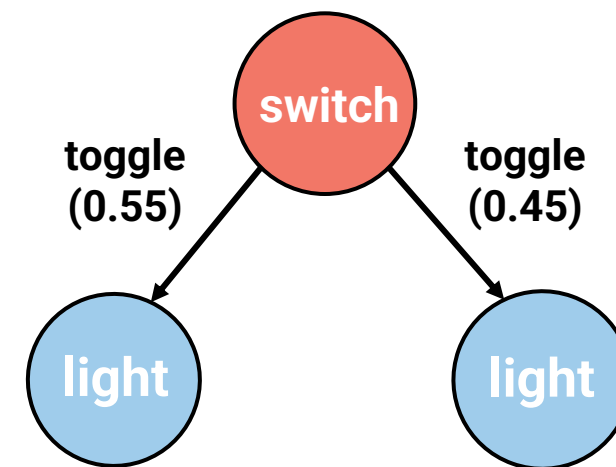
Encode edges as binary variables  $E_1, E_2$

## Factor Graph (FG)



Construct a preliminary FSG from visual input

## Probabilistic FSG



Joint factor graph inference yields per-edge confidence

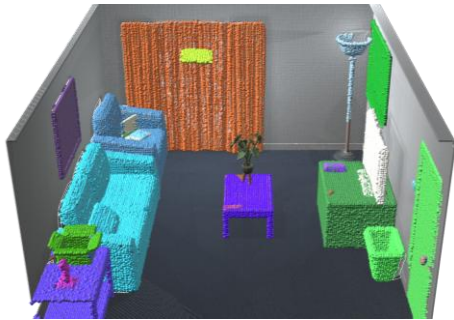
# Method Overview

FunFact builds **probabilistic, open-vocabulary** functional 3D scene graphs from posed RGB-D images.

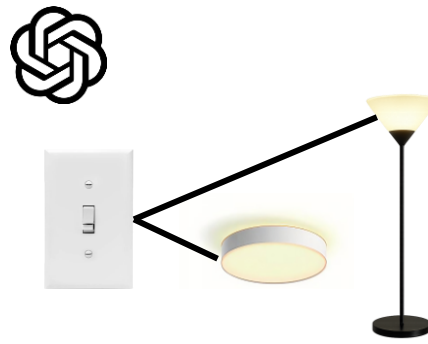
Posed RGB-D



3D Scene Reconstruction

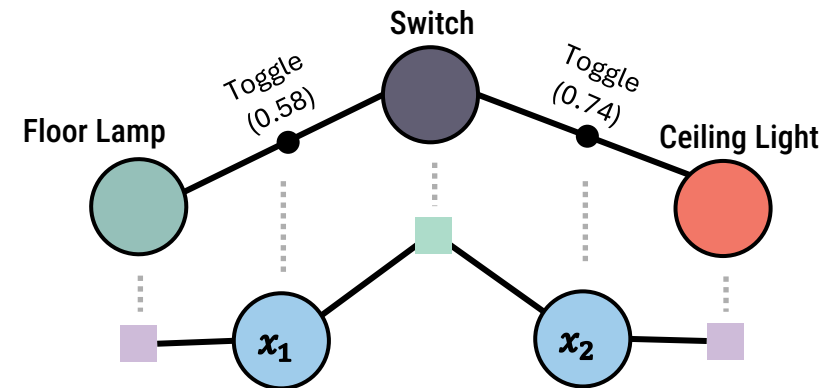


Functional Relation Template



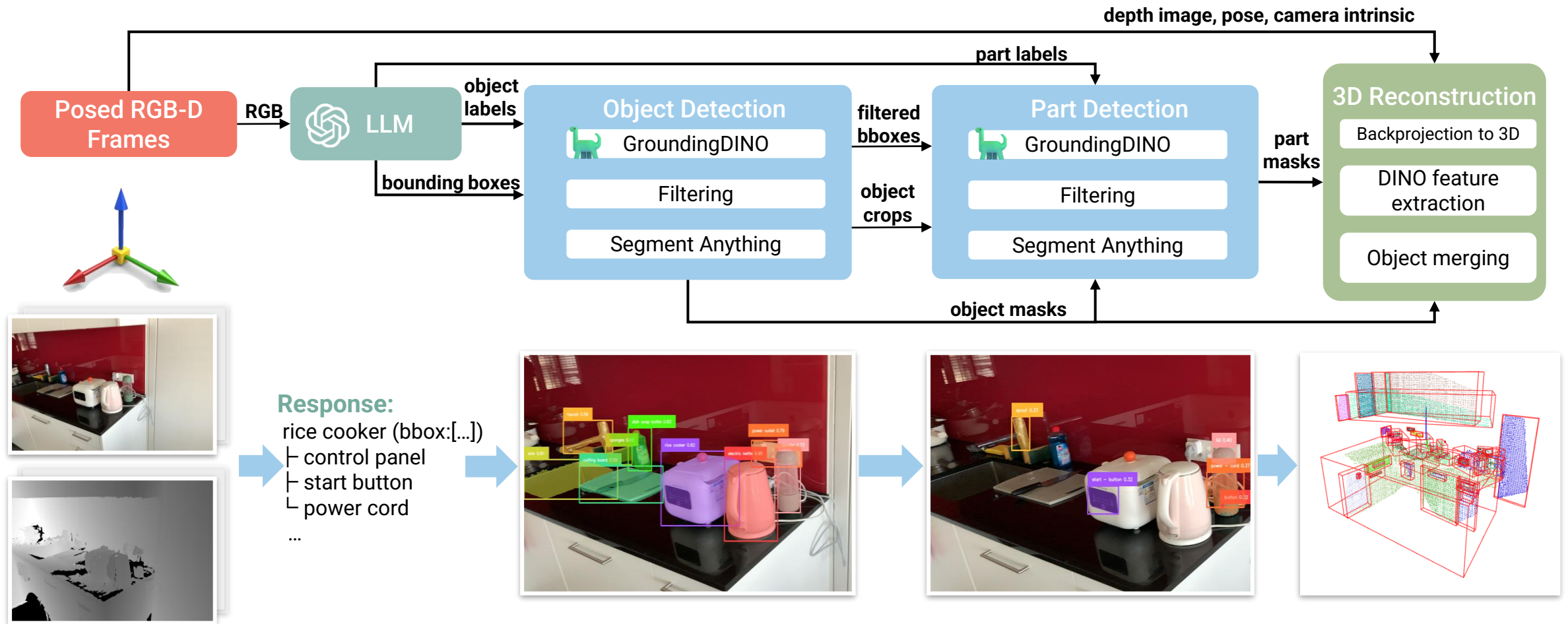
— toggle (one-to-one)

FG Construction & Probabilistic FSG

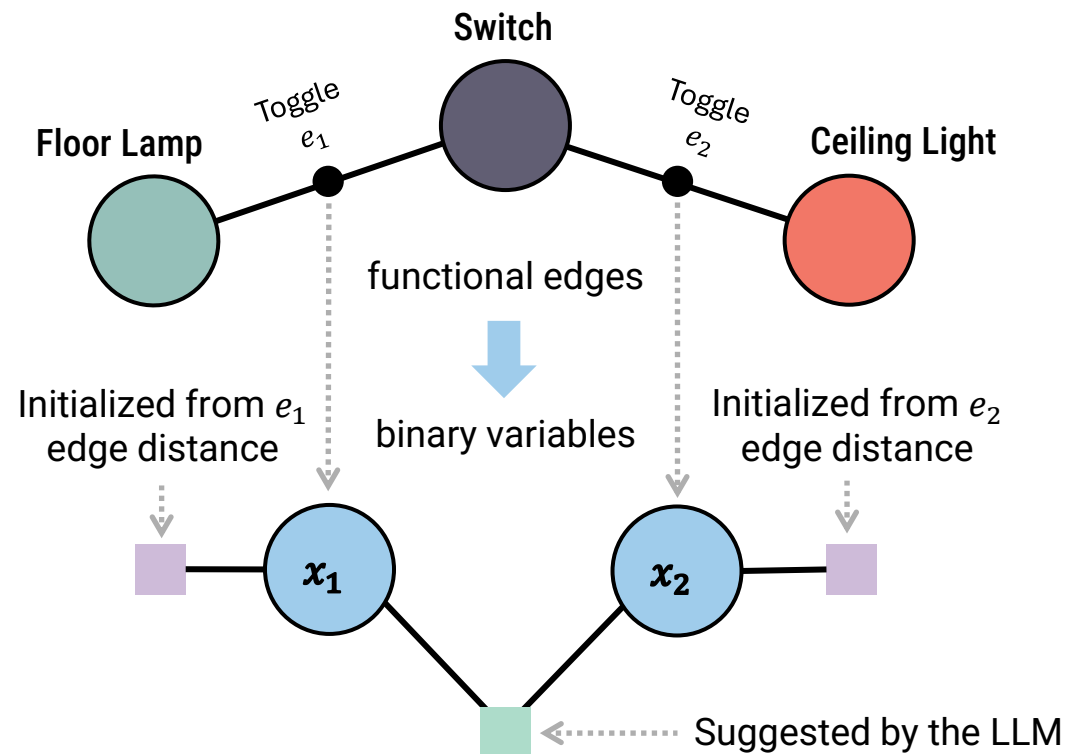


Factor Variable

# Method Overview >3D Scene Reconstruction



# Method Overview >Dual Factor Graph Construction



## Proximity prior

For edge  $e_i$  with the Euclidean distance  $d(e_i)$  and its dual variable  $x_i$ :

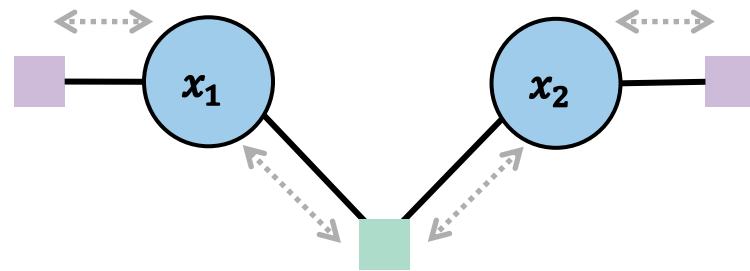
$$\phi_{prox}(x_i) = \exp\left(-\frac{d(e_i)}{\lambda}\right)$$


## Cardinality prior

For a part node  $v$ , let  $X_v$  denote the incident variables, and let  $d_n = \sum_{x \in X_v} x$ ,  $b \in (0,1)$ :

$$\phi_{card}(X_v) = \begin{cases} b^{d_n-1} & \text{if } d_n \geq 1, \\ b^2 & \text{if } d_n = 0, \end{cases}$$

# Method Overview > Probabilistic Inference



 Message passing  
(loopy belief propagation)



## Joint Inference

Run loopy propagation on the dual factor graph to reason over all functional edges jointly

## Marginal Probabilities

For each edge variable  $x_i$ , compute its marginal probability:

$$P(x_i = 1 \mid \text{all context})$$

## Edge Confidence

These marginals become per-edge confidence scores that reflect both local ambiguity and global scene context.

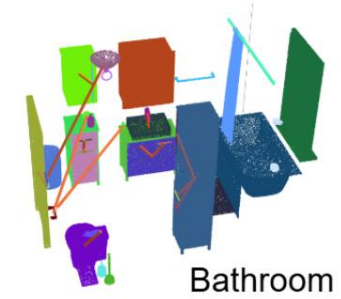
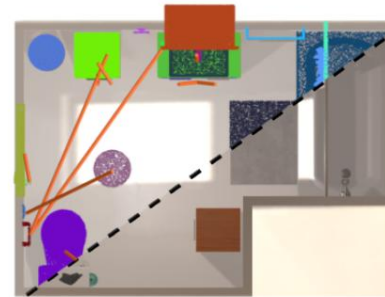
# Evaluation on FunGraph3D

Methods	Objects (↑)		Interactive Elements (↑)		Functional Edges (↑)	
	Recall@3	Recall@10	Recall@3	Recall@10	Recall@5	Recall@10
OpenFunGraph	70.7	79.1	44.4	57.6	29.8	45.0
<b>FunFact (Ours)</b>	<b>91.1</b>	<b>96.6</b>	<b>68.3</b>	<b>78.7</b>	<b>48.7</b>	<b>63.9</b>

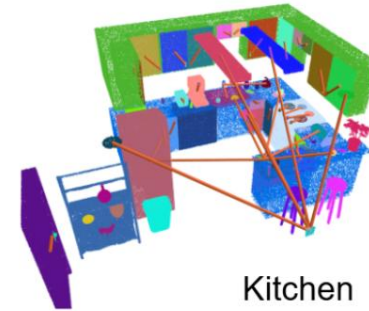
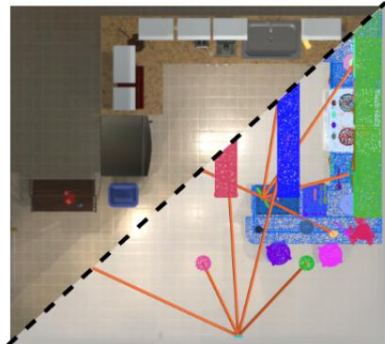
# New Dataset - FunThor



AI2-THOR – An interactive Embodied AI Simulator



Bathroom



Kitchen

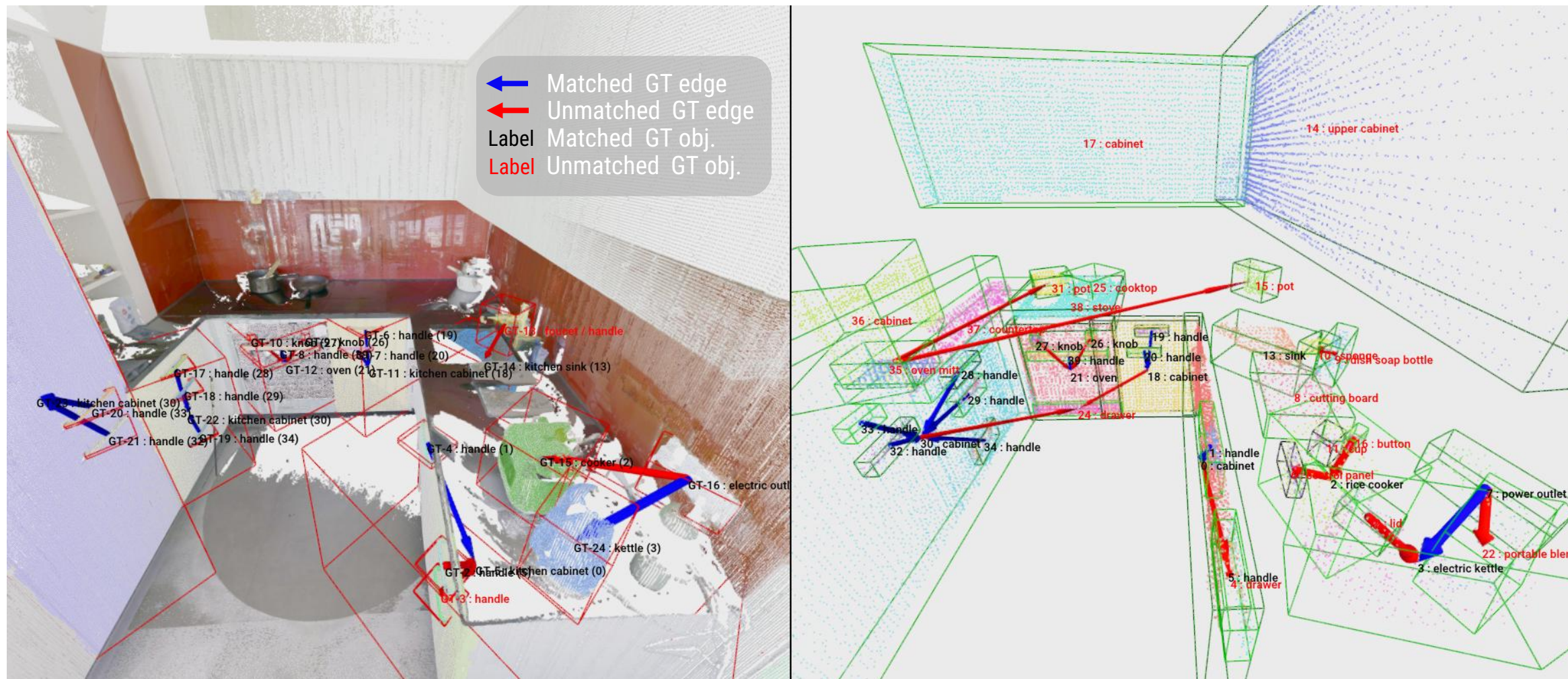
FunThor – An automatically annotated dataset based on AI2-THOR

# Comprehensive Evaluation on FunThor

Method	Prec. (↑)	Recall (↑)	F1 (↑)	ECE (↓)*
OpenFunGraph	23.4	12.2	16.0	0.25
<b>FunFact (Ours)</b>	<b>31.9</b>	<b>49.3</b>	<b>38.7</b>	<b>0.11</b>

\*ECE: Expected Calibration error. Lower is better.

# Qualitative Example



Ground Truth

Reconstruction & Prediction

Thank you!



**Project Page**



Scan to try the  
*interactive* results!

[funfact-scenegraph.github.io](https://funfact-scenegraph.github.io)