SFU SIMON FRASER UNIVERSITY

When Language Meets the Physical World: Semantic Representation for Robots

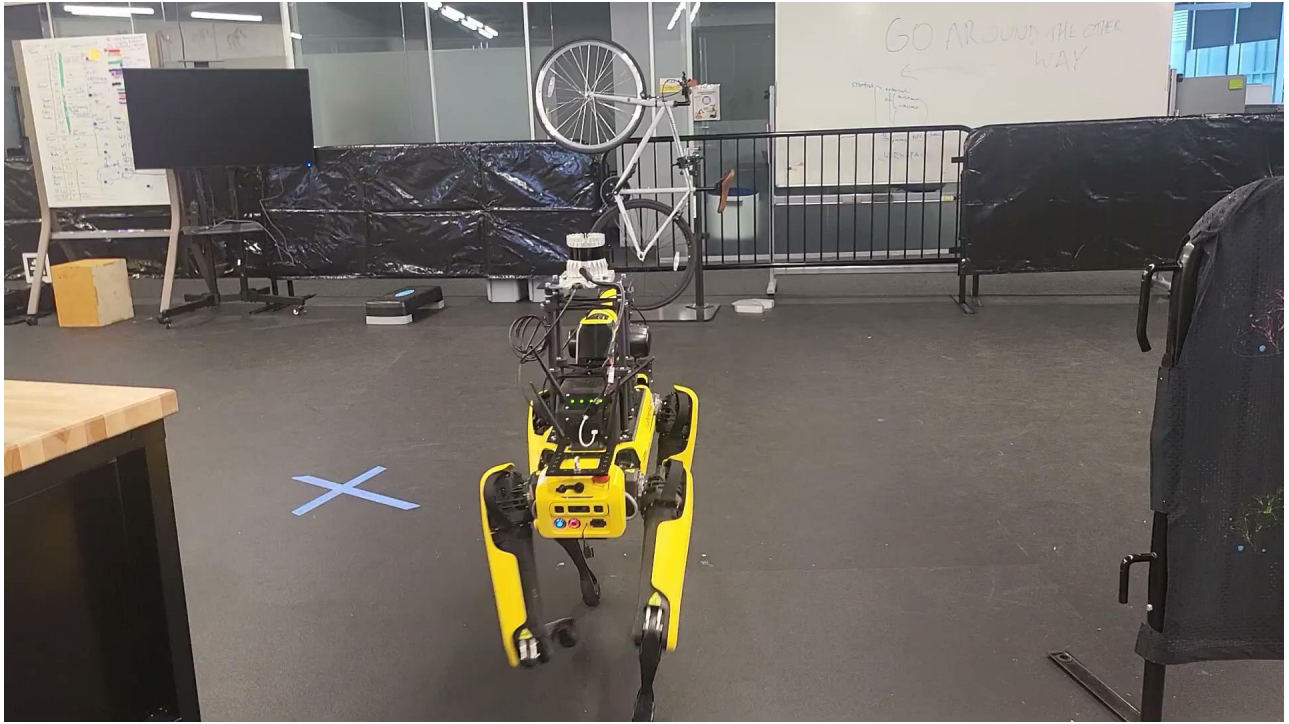
Presenter: Sonia Raychaudhuri
March 30, 2026


TESLA
TESLA
TESLA


PERSONAL AI ROBOTS PET

SFU SIMON FRASER UNIVERSITY

Move forward to the bicycle. Turn right, then move to the chair. Turn left, and stop near the potted plant.



Raychaudhuri et al. Zero-shot Object-Centric Instruction Following: Integrating Foundation Models with Traditional Navigation. 2025

SFU SIMON FRASER UNIVERSITY

Language must be grounded in perception

Move forward to the bicycle. Turn right, then move to the chair. Turn left, and stop near the potted plant.

- What does *move forward* mean?
- What is a *bicycle*?
- How does it look?

— Robots need grounding

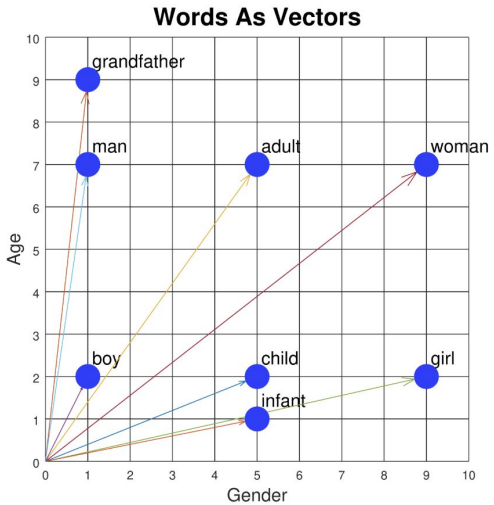
Raychaudhuri et al. Zero-shot Object-Centric Instruction Following: Integrating Foundation Models with Traditional Navigation. 2025

SFU SIMON FRASER UNIVERSITY

From NLP to Robotics

NLP:

Text → Embeddings → Meaning



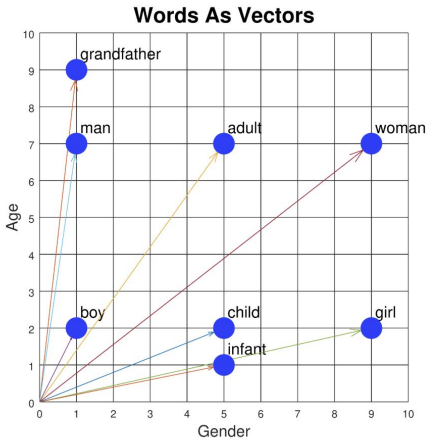
https://www.cs.cmu.edu/~dst/WordEmbeddingDemo/tutorial.html

SFU SIMON FRASER UNIVERSITY

From NLP to Robotics

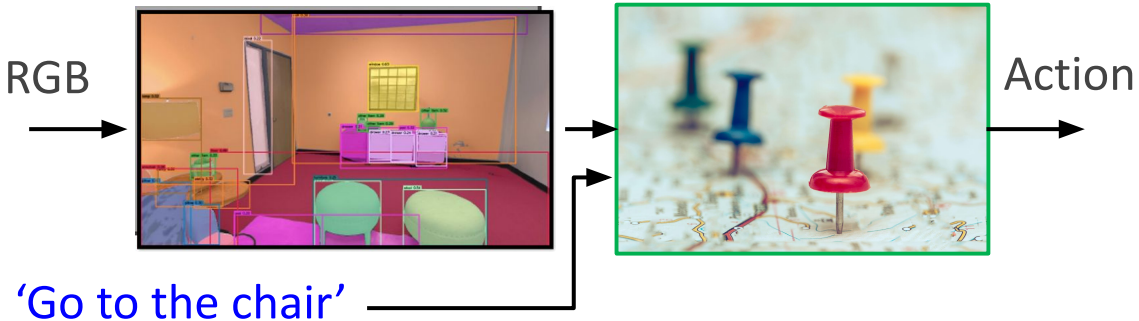
NLP:

Text → Embeddings → Meaning



Robotics:

Images → Perception →
Language → Embedding →
Semantic Map → Action

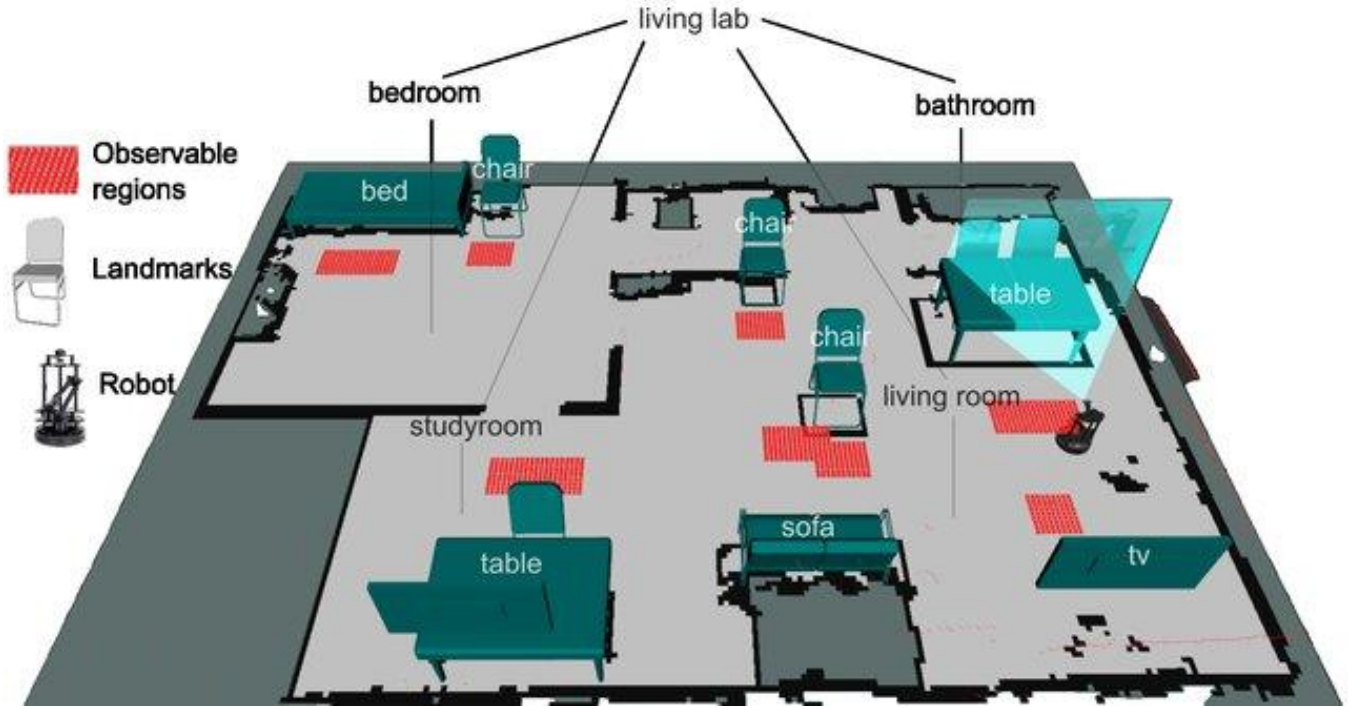


https://www.cs.cmu.edu/~dst/WordEmbeddingDemo/tutorial.html

Gu et al. ConceptGraphs: Open-Vocabulary 3D Scene Graphs for Perception and Planning. 2023.

SFU SIMON FRASER UNIVERSITY

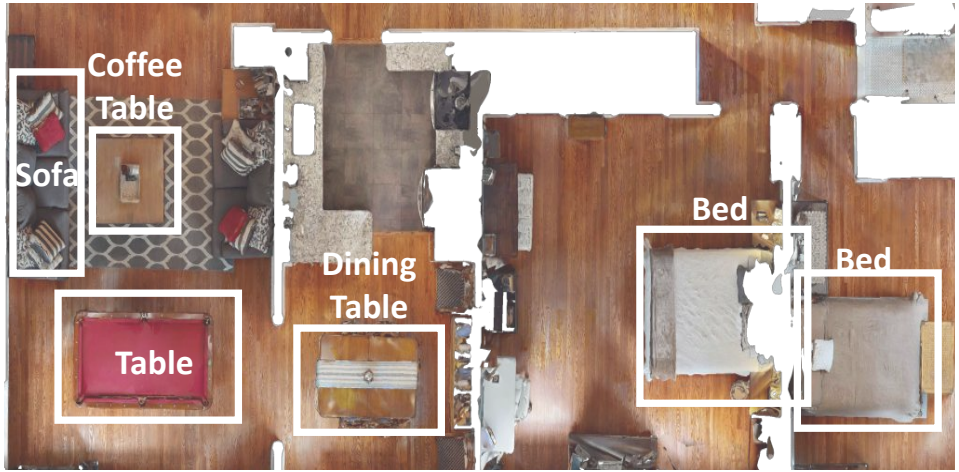
What is a Semantic Map?



Semantic-Based Multi-Object Search Optimization in Service Robots Using Probabilistic and Contextual Priors. 2024

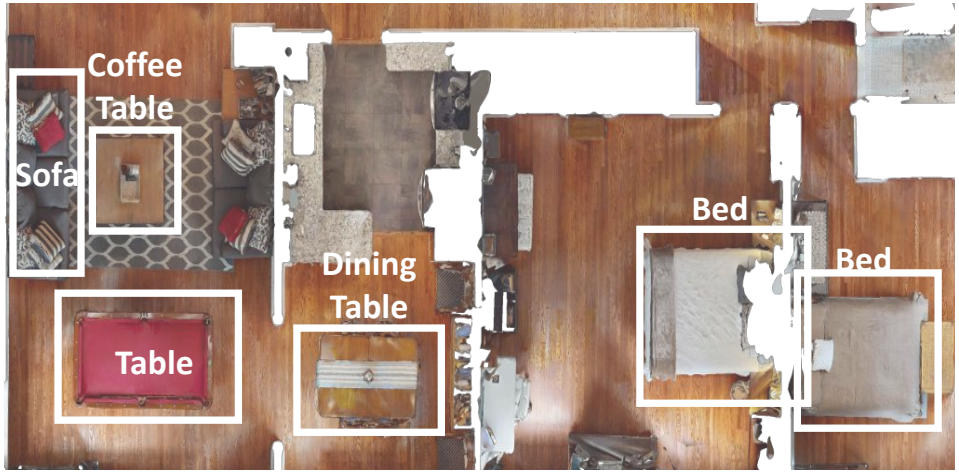
SFU SIMON FRASER UNIVERSITY

Map Structure

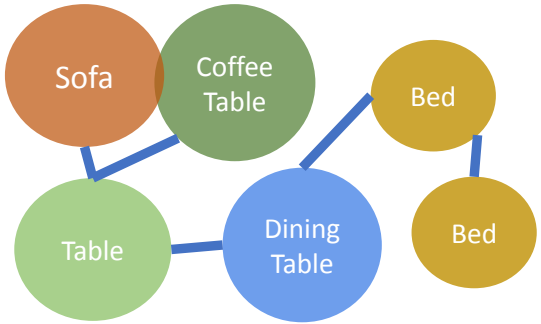


SFU SIMON FRASER UNIVERSITY

Map Structure

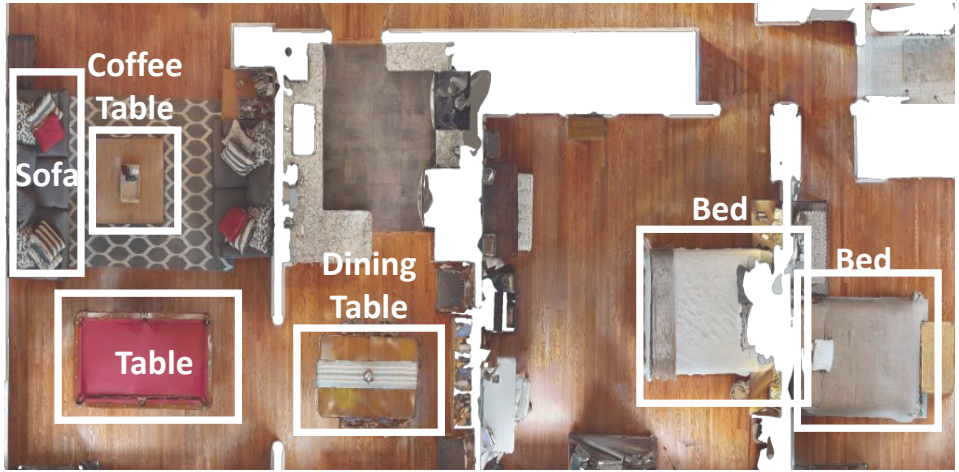


Topological Map

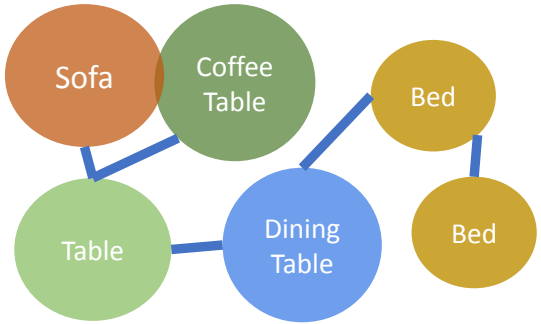


SFU SIMON FRASER UNIVERSITY

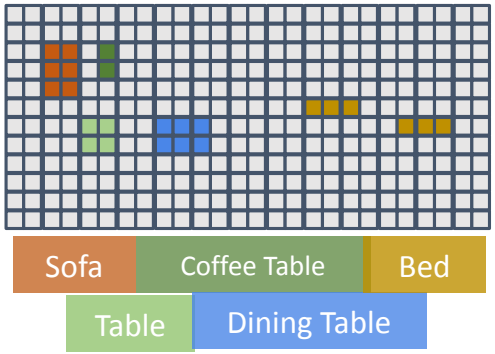
Map Structure



Topological Map

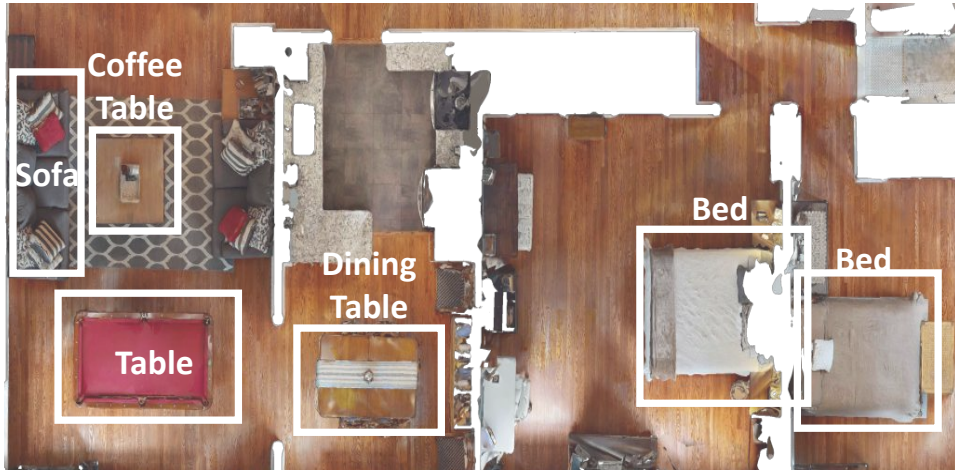


Grid Map

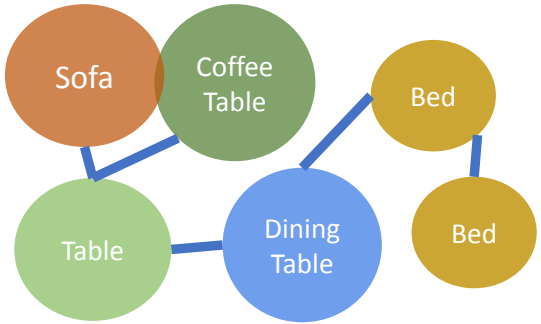


SFU SIMON FRASER UNIVERSITY

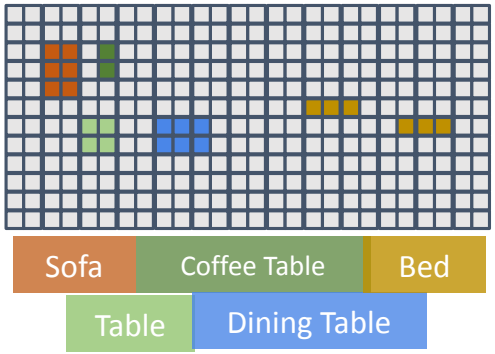
Map Structure



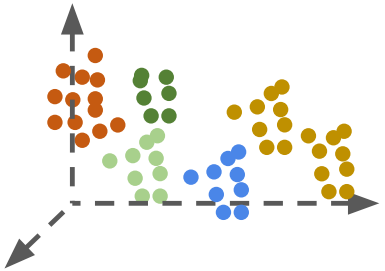
Topological Map



Grid Map

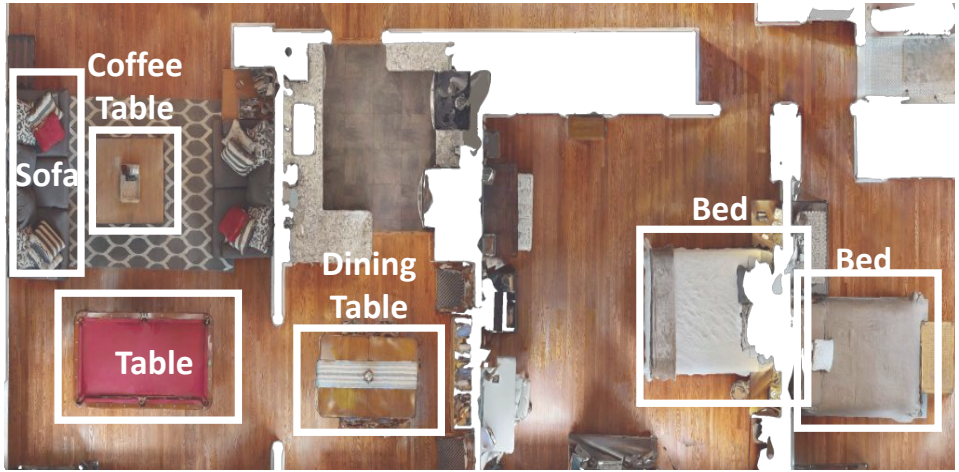


Geometric Map

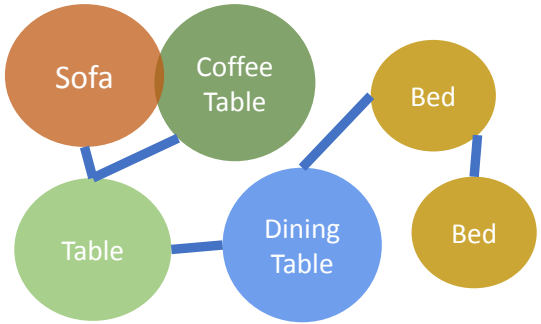


SFU SIMON FRASER UNIVERSITY

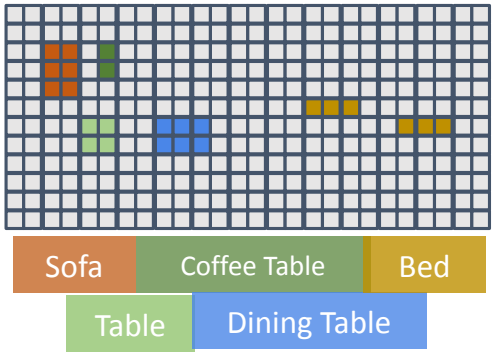
Map Structure



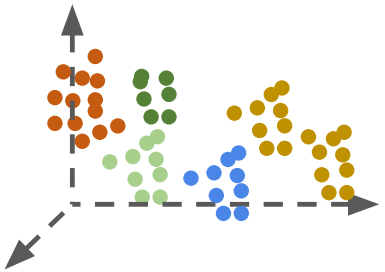
Topological Map



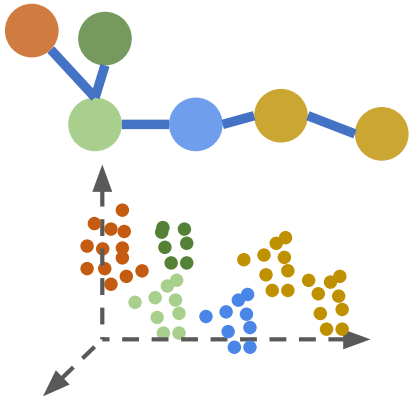
Grid Map



Geometric Map

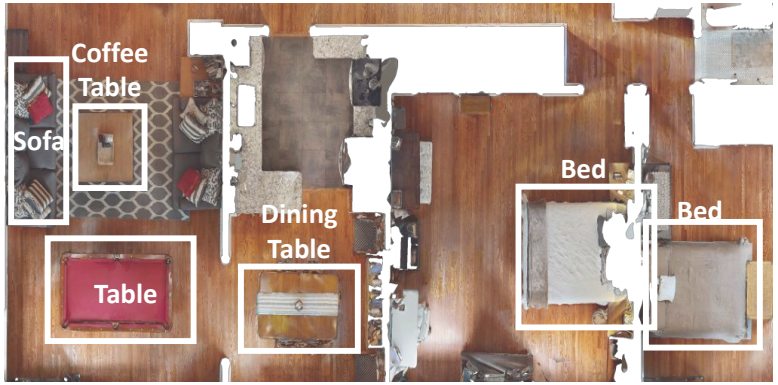


Hybrid



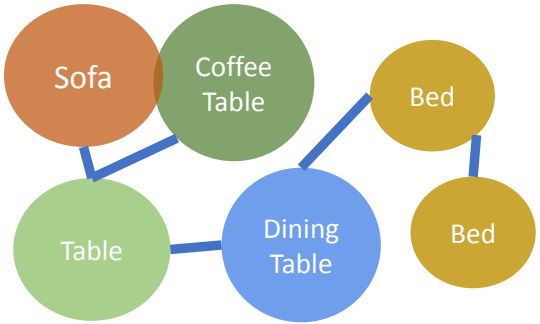
SFU SIMON FRASER UNIVERSITY

Map Structure

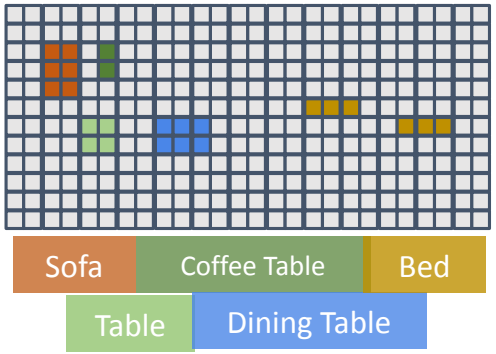


Sparse ⟷ Dense

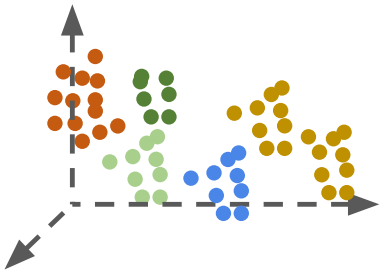
Topological Map



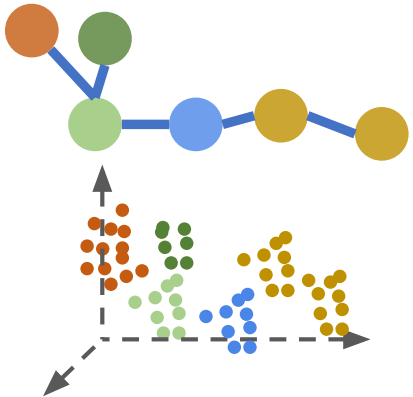
Grid Map



Geometric Map

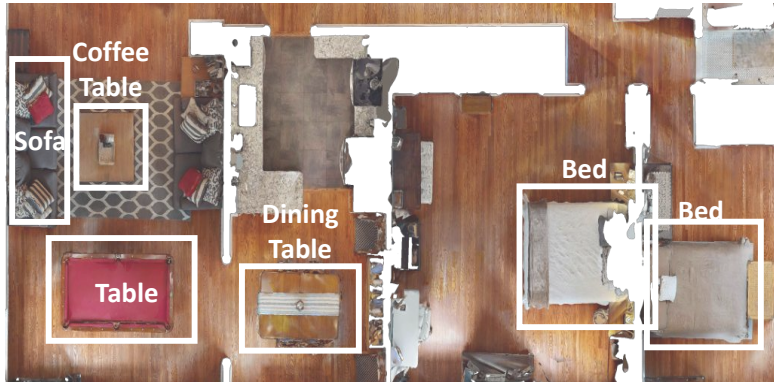


Hybrid



SFU SIMON FRASER UNIVERSITY

Map Encoding

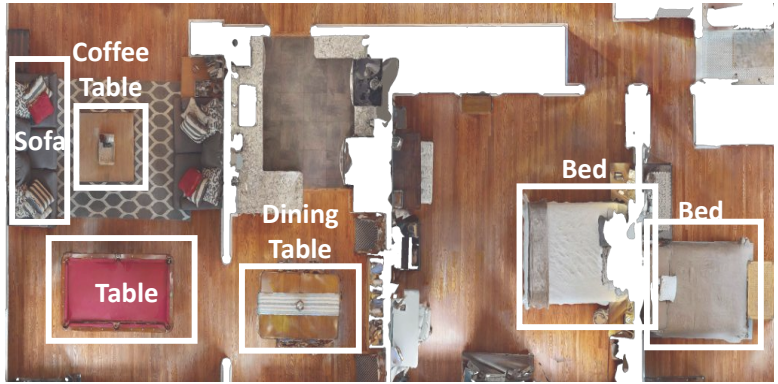


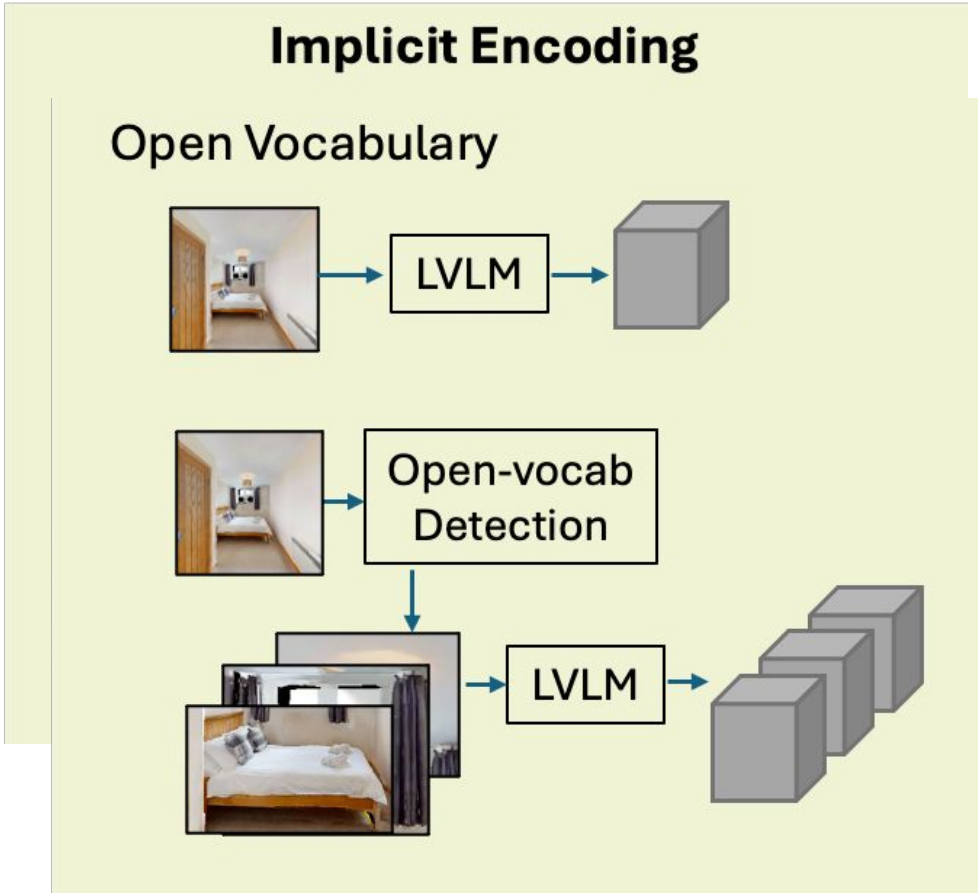
Explicit Encoding

Occupancy	Binary	[0,1]
Image	RGB pixel values	
Objects	RGB pixel values	
	One-hot for n predefined object types	
	Text labels for object types	“bed”
	room types	
	Text labels for room types	“bedroom”
Audio Intensity	Single value in the range	[0-1]

SFU SIMON FRASER UNIVERSITY

Map Encoding





SFU SIMON FRASER UNIVERSITY

How It Enables Robot Tasks

goal becomes meaningful, not just geometric

Navigation
“go to kitchen”

Object search
“find mug”

Interaction
“pick up object”

Without semantics

- random exploration
- inefficient search

With semantics

- directed navigation
- intelligent search

SFU SIMON FRASER UNIVERSITY

Key Challenges

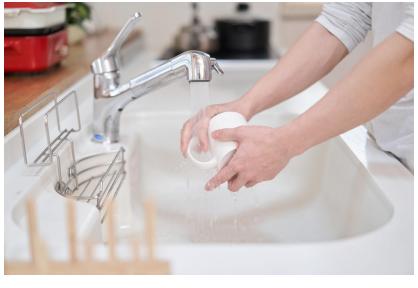
Why is this still an open research problem?

Language is Ambiguous

'go to the orange scooter'

Dynamic Environments

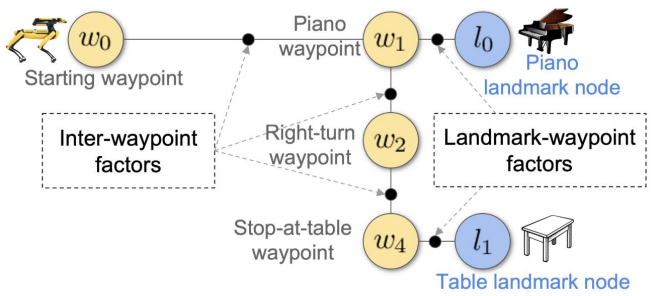
mug on the table → mug in sink



Generalization

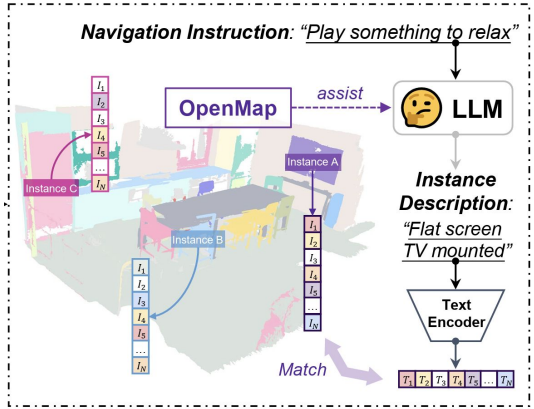
Future Directions

Language as Knowledge (LLMs for Robots)



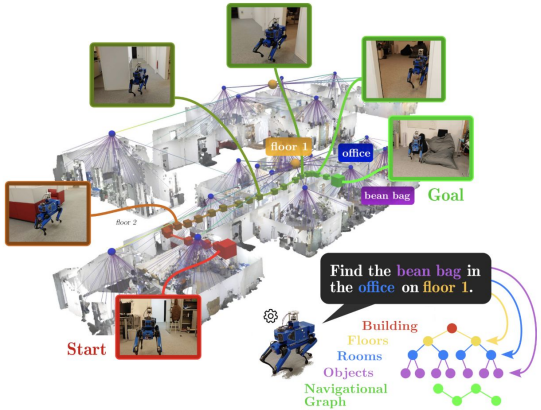
Raychaudhuri et al. “Zero-shot Object-Centric Instruction Following: Integrating Foundation Models with Traditional Navigation.” [2025]

Open-World Understanding



Li et al. "OpenMap: Instruction Grounding via Open-Vocabulary Visual-Language Mapping." [2025]

Hierarchical Representations



Werby et al. “Hierarchical Open-Vocabulary 3D Scene Graphs for Language-Grounded Robot Navigation.” [2024]

SFU SIMON FRASER UNIVERSITY

Takeaways

The next step for NLP is not just understanding text — but understanding the world

and semantic representations are the bridge between language and physical world