# A Survey of Robotic Language Grounding: Tradeoffs between Symbols and Embeddings

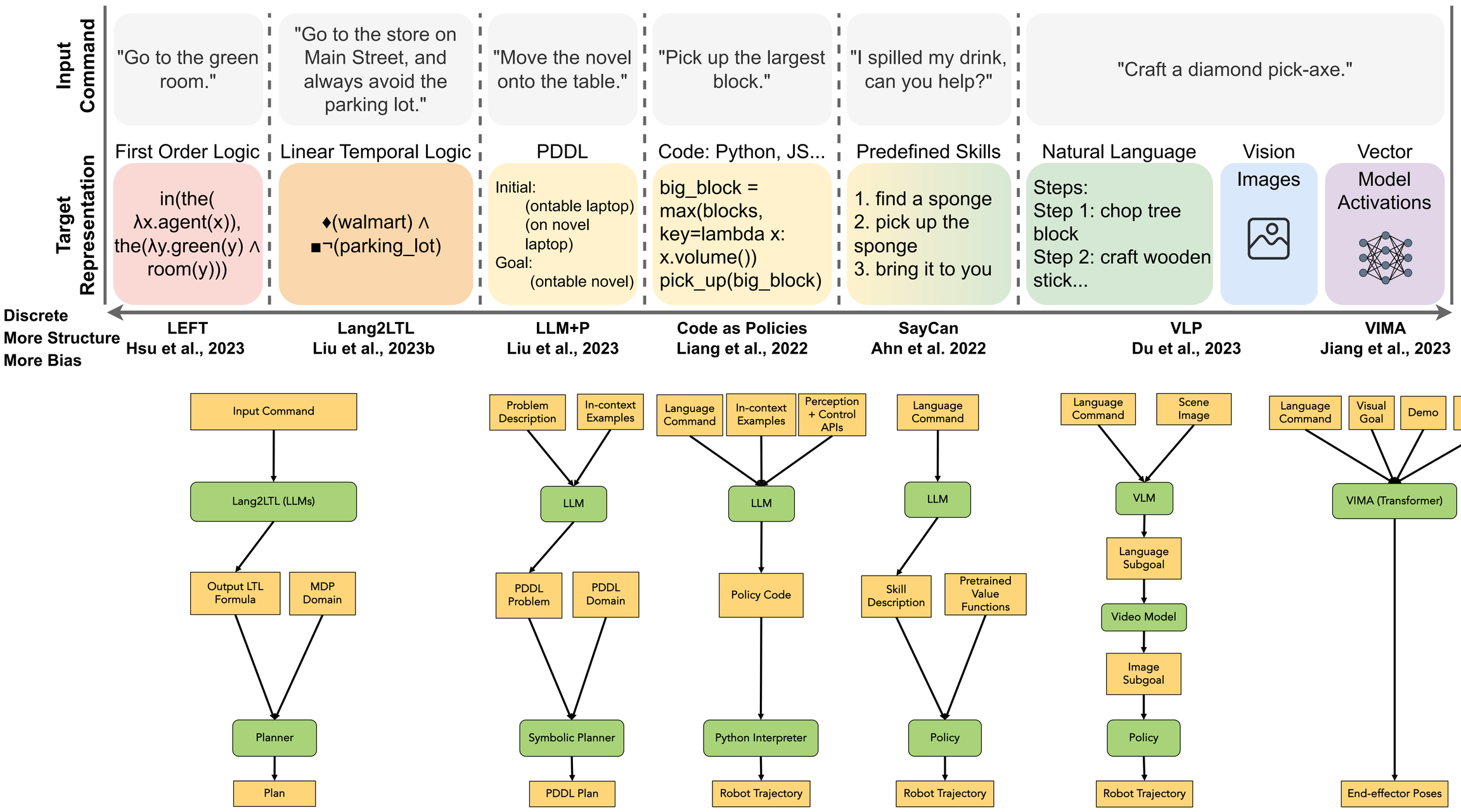
Vanya Cohen (vanya@utexas edu), **Jason Xinyu Liu** (xinyu_liu@brown.edu),
Raymond Mooney, Stefanie Tellex, and David Watkins

## Robotic Language Grounding

Connect linguistic elements in language to the robot's perception of and actions in the physical world.

- What grounding representation to use?
- How to ground natural language to the grounding representation of choice?

## A Spectrum from Symbols to High-dimensional Embeddings




## Why This Spectrum

### Symbols: Logic, PDDL

- Discrete; More Structure; More Bias

✓ Unambiguous semantic
✓ Verifiable
✓ Interpretable
✓ Reduce search space
✓ Sound, complete, (often) optimal

✕ Require manually defined structures
✕ Difficult to represent low-level control

### Symbols: Code, Predefined Skill Descriptions

✓ Flexible
✓ Adaptive
✓ LLMs good at general purpose code writing

✕ Required predefined perception and control models
✕ May generate incorrect plans

### High-dimensional Embeddings

- Continuous; Less Structure; More variance

✓ Adaptive

✕ May generate incorrect plans

## Open Problems and Future Directions

### Neuro-symbolic Approach

- POMDP and PDDL planners
- Deep learning models with generalizable representations
- E.g., Jointly learn symbols in the embedding space and skills

### Multimodal Dataset

- E.g., text, audio, RGB images, point clouds, voxels, videos, demonstrations
- Semantically diverse

### Modular Approach

- Existing robot modules
- E.g., SLAM, motion planning and object detection

### Verification and Safety

- Formal methods

https://arxiv.org/abs/2405.13245

IJCAI JEJU 2024
International Joint Conference on Artificial Intelligence
Jeju 03.08.24 - 09.08.24

Paper 8081

https://jasonxyliu.github.io