



Enabling Robots to Understand Language

Presentation at
IIT Bombay NLP Course
October 17th, 2025

Prof. Rohan Paul

Associate Professor, Dept. of Computer Science and Engineering
Joint Appt., School of Artificial Intelligence and Associate Faculty, School of IT
Indian Institute of Technology Delhi

Talk Outline

- Intelligent robots and the need to understand language
- Challenges in understanding language
- Research efforts and key results [2015 - present]
- Linkages between language understanding – perception – action
- Outstanding questions

The Evolution of Machines in Society

Agricultural Age

Plough



Industrial Age

Engine



Electronics Age

Computers



Robotic age

Sensing & compute



Increased capacity to sense
interact and manipulate the
physical world

Future Robots: Highly capable and intelligent

Cognitive Intelligence

Past

Robot as mechanical tools



Continuous control by human.
E.g., DaVinci

Present

Remote controlled robots



- Controlled by a remote operator
- Some ability to move on its own.
- E.g., PackBot Scout robot.

Future: AI-based Cognition

Intelligent Robotic Team mate



“Robot, examine if the crate is empty”.
No, I can see a grenade.



- Controlled via verbal commands (in many languages)
- Autonomous High-level task execution
- ChatGPT powered autonomous reasoning.
- Work with humans with intermittent guidance.
- Platform: wheels, track, legged, air borne.

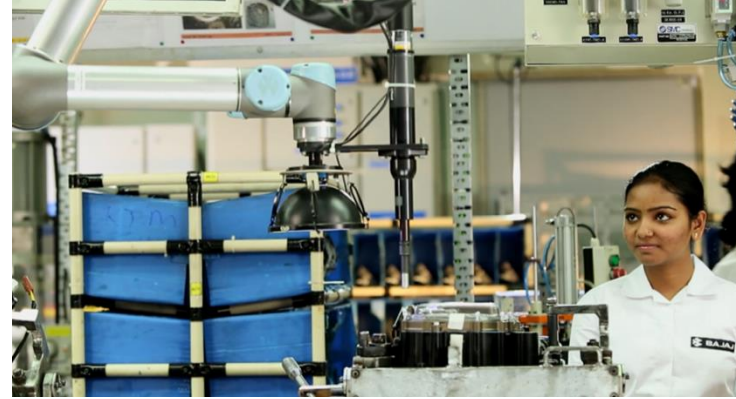
Evolution of Robots

Language: An intuitive means to communicate

“Prepare my breakfast”



“इंजन भागों को इकट्ठा करें”



“अटैची का निरीक्षण करें”



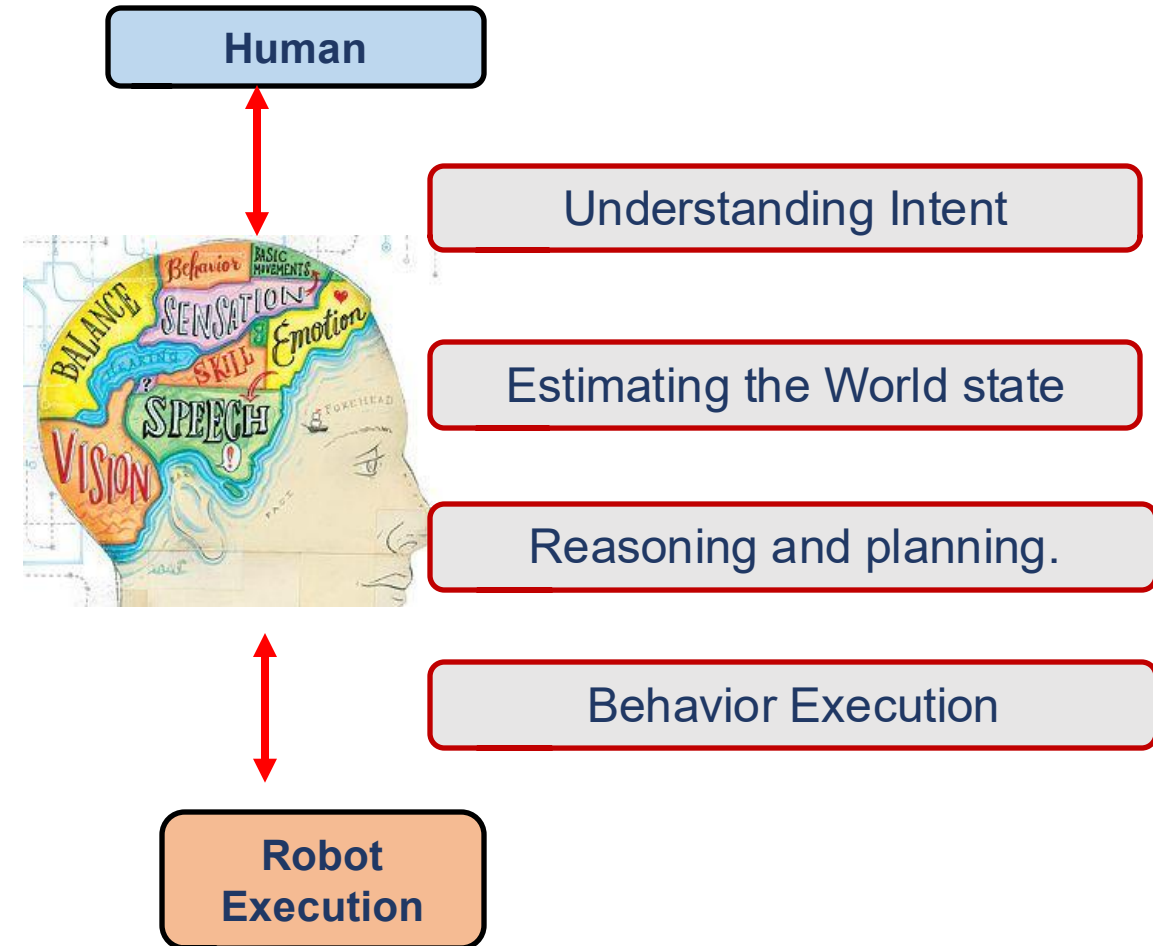
“Go to the fire hydrant behind the traffic cone”



How to infer goals from language instructions?

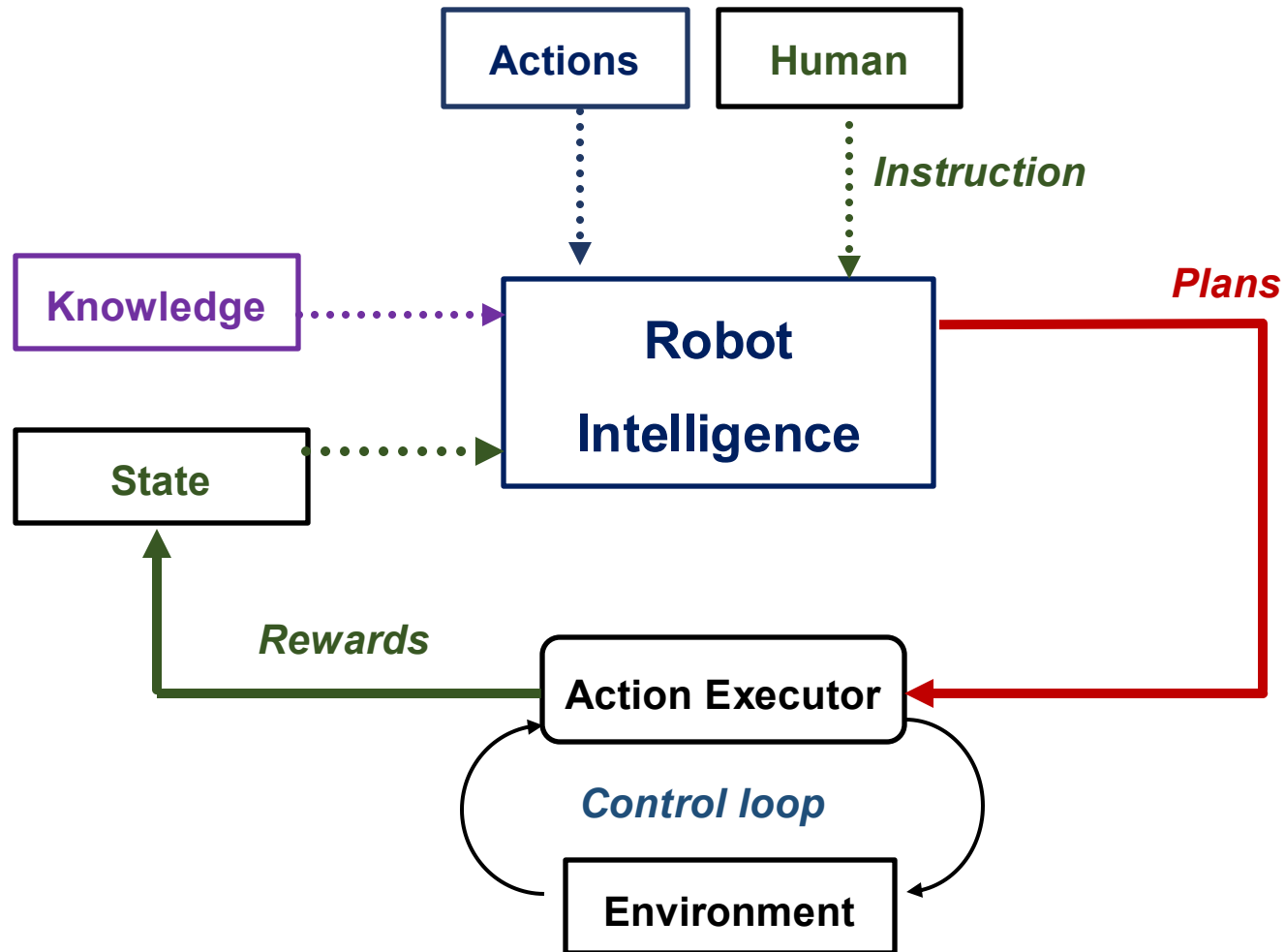
High-level Intelligence

Robot, Is there an object on the ground? Clear the objects on the ground. Check if the box is empty.



**A robot needs to transform the world state as per the human's instruction.
High goal reaching rate. Robustness under uncertainty.**

Robot as a goal-directed agent



Estimation:

$F(\text{Actions}, \text{Instruction}, \text{Knowledge}, \text{State}) \rightarrow \text{Plans}$

Desired:

- High-goal reachability
- Diversity goals - Complex multi-steps
- Interpretable (communicable)

Model

- View as an MDP (diverse range of goals)
- Goal—conditioned policies.
- Complex state with constraints

What makes language understanding hard?

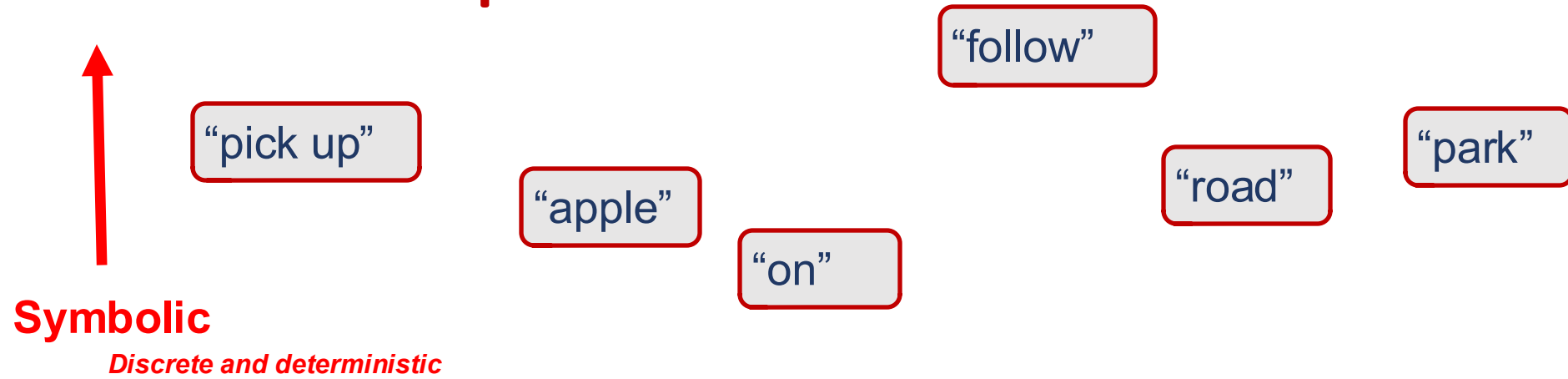
“Pick up the apple on the tray”



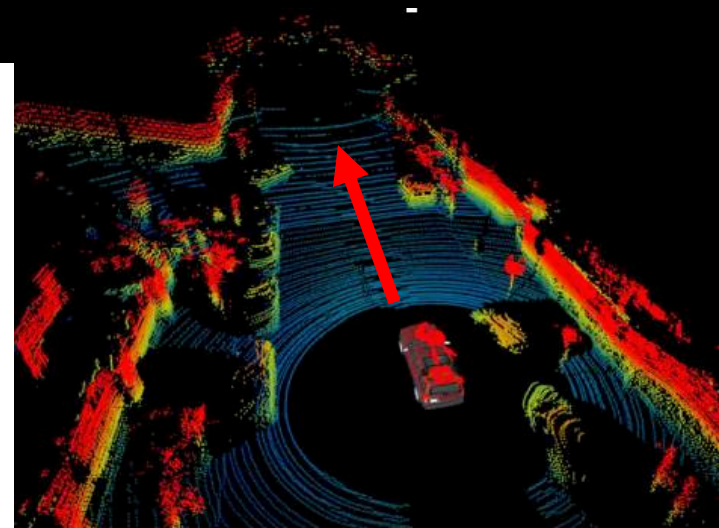
“Follow the road to the parks”



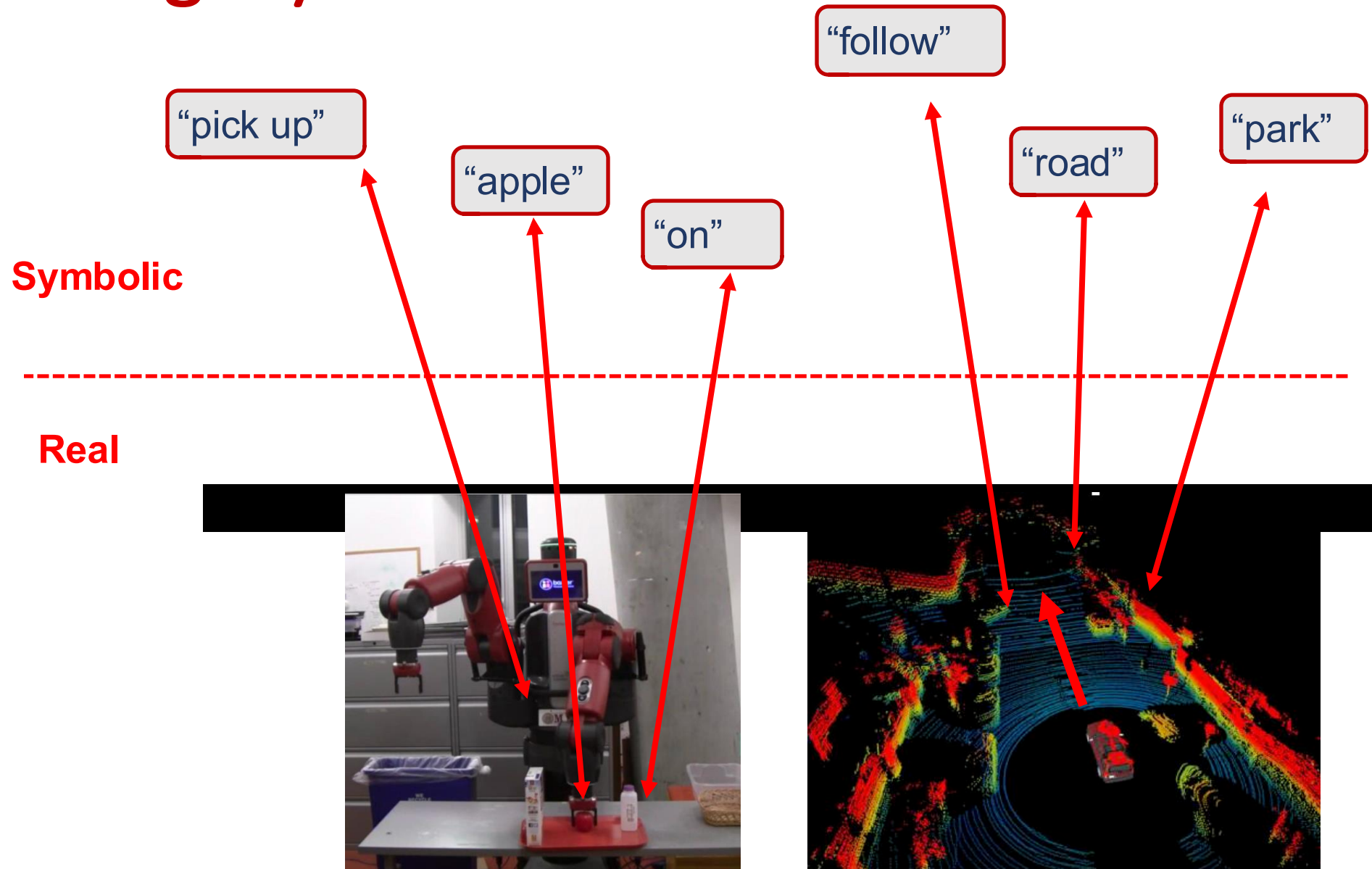
The Semantic Gap



Real
Continuous and noisy

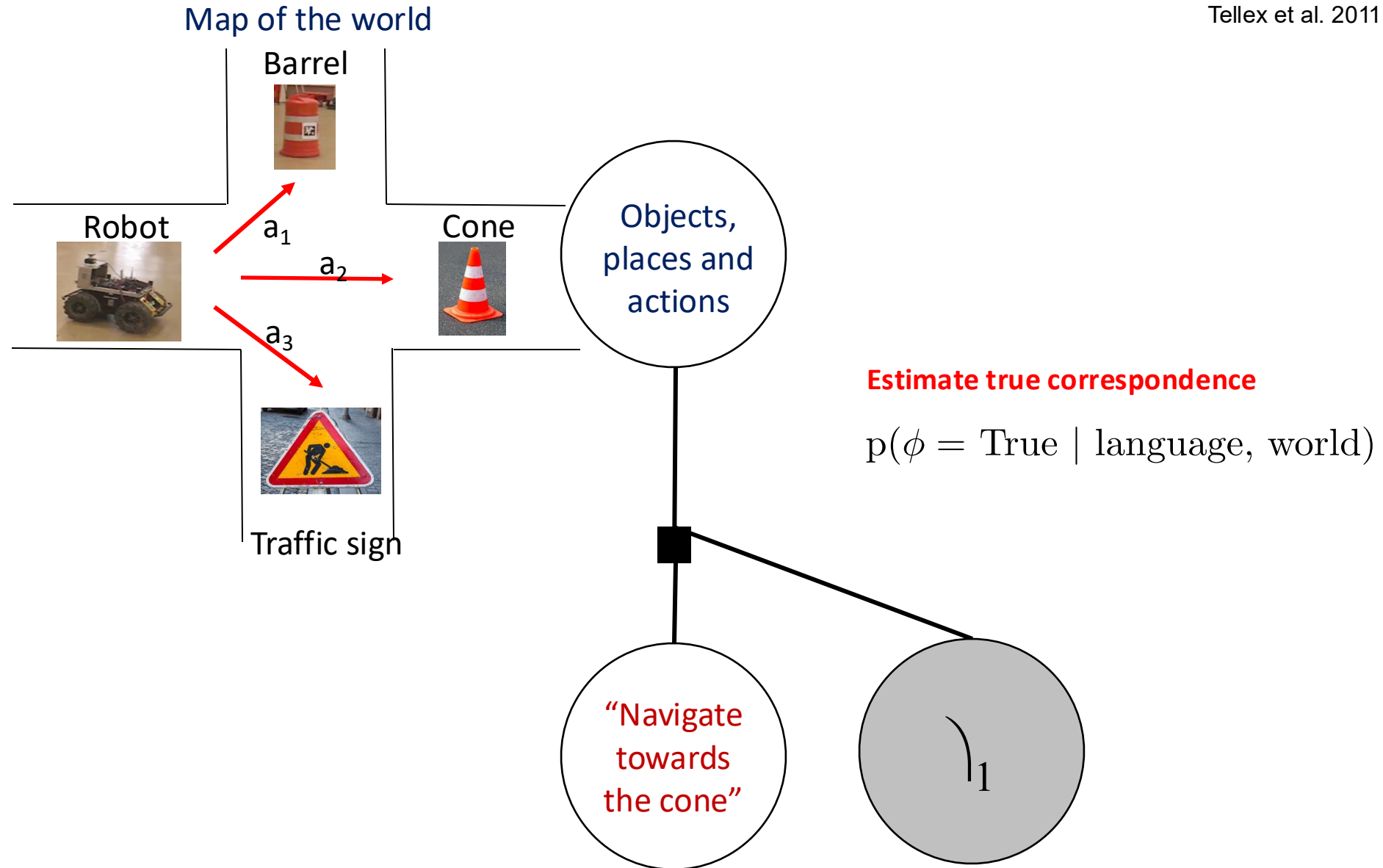


Grounding “symbolic” to “real”



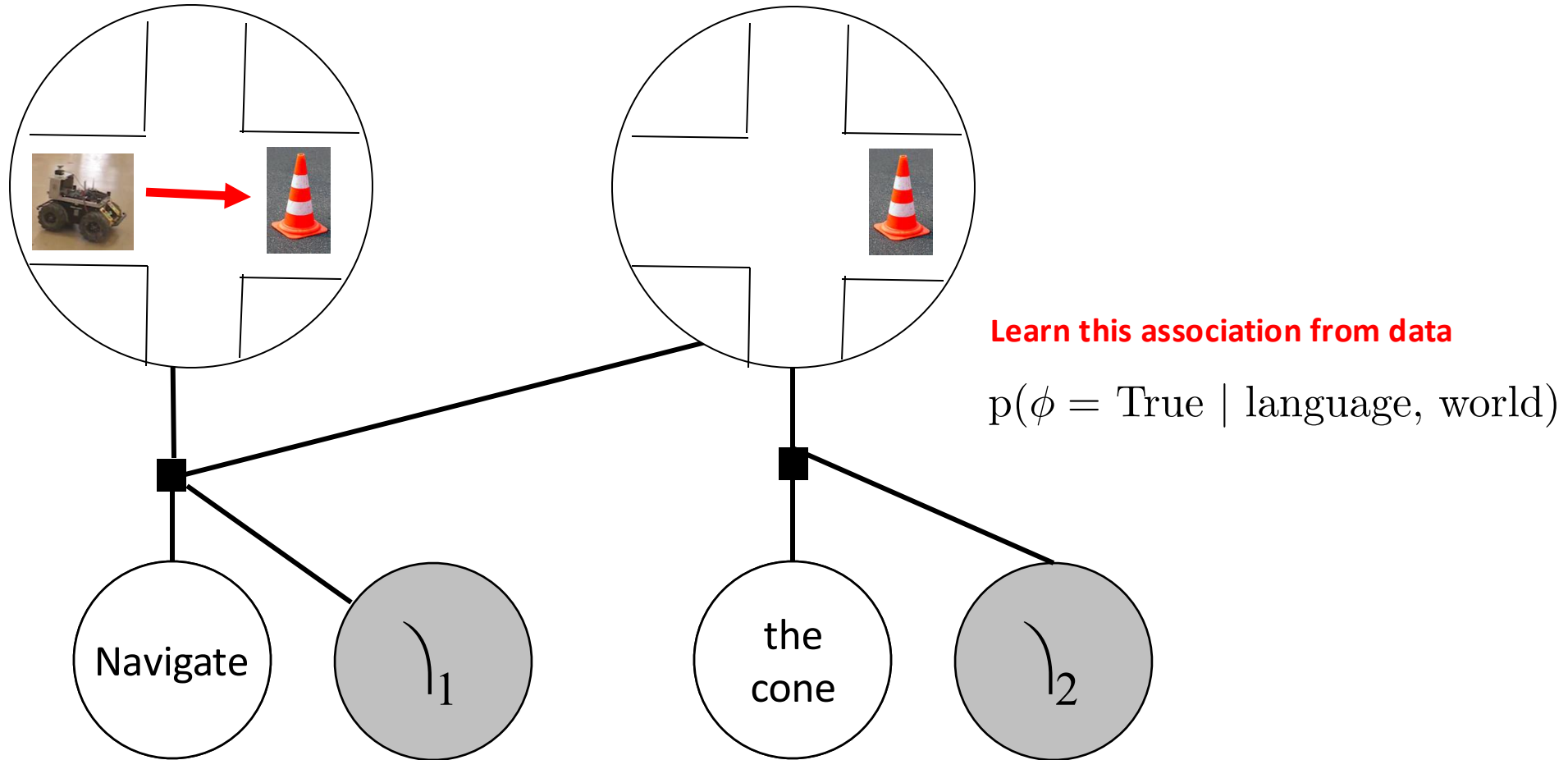
Generalized Grounding Graphs

Tellex et al. 2011



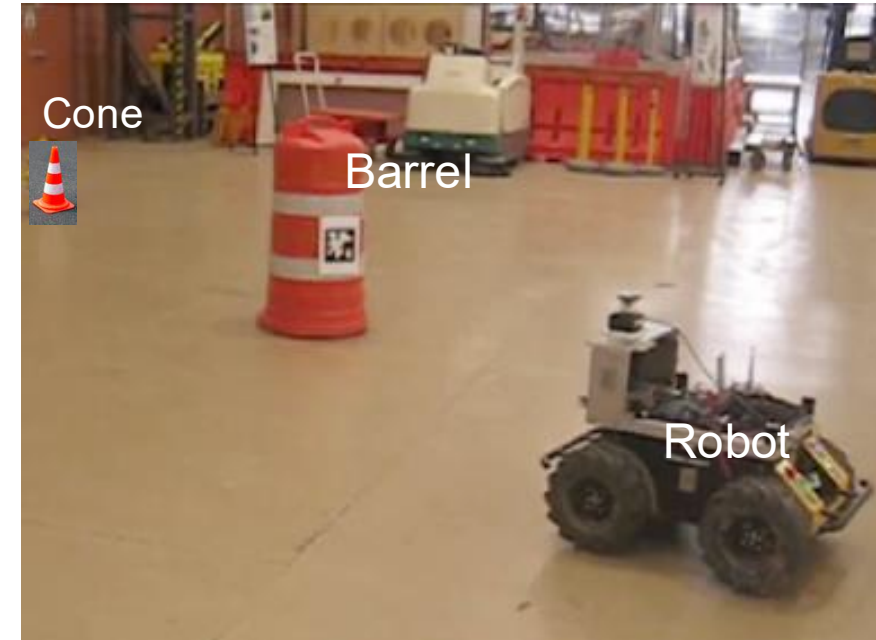
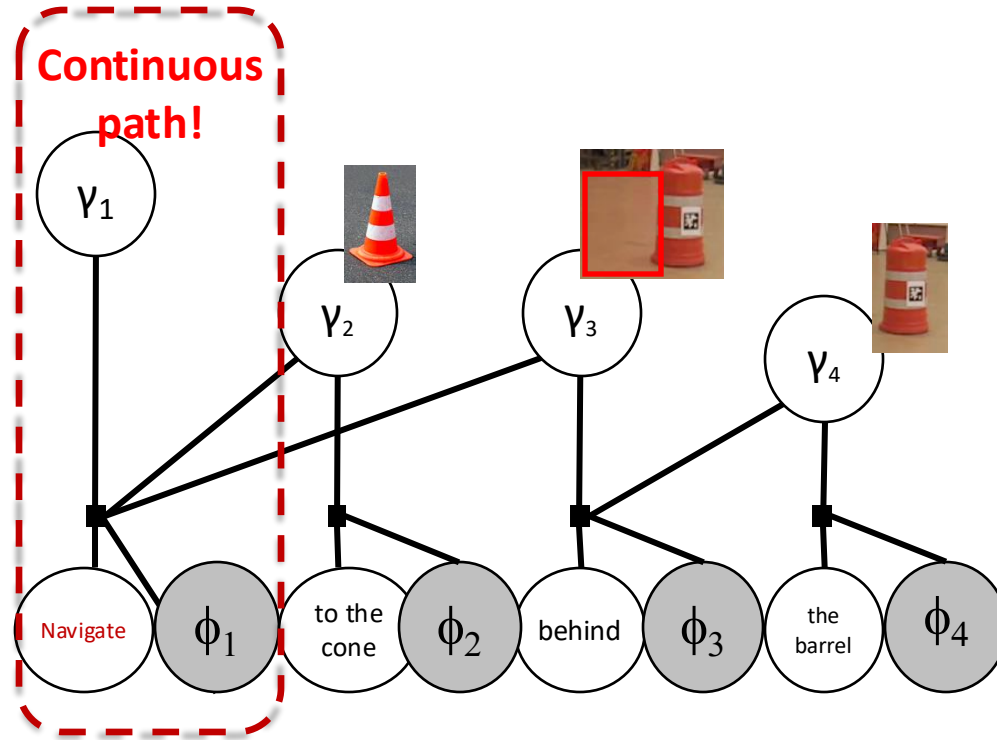
Generalized Grounding Graphs

Tellex et al. 2011



Interpreting an Instruction

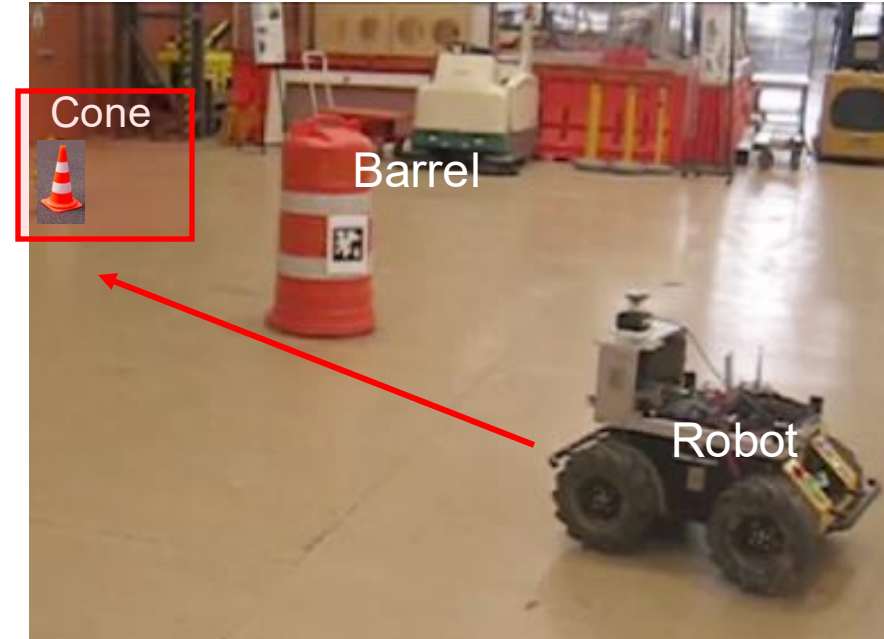
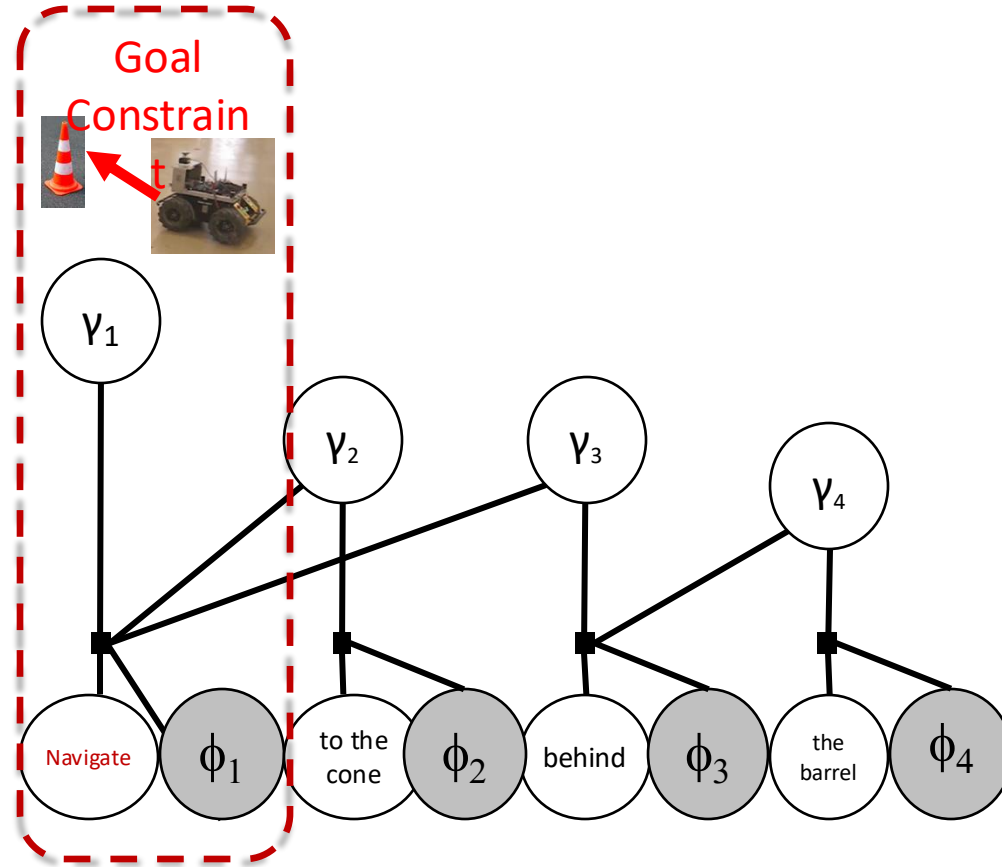
Howard et al. 2014



“Navigate to the cone behind the barrel”

Interpreting an Instruction: Constraints

Howard et al. 2014



“Navigate to the cone behind the barrel”

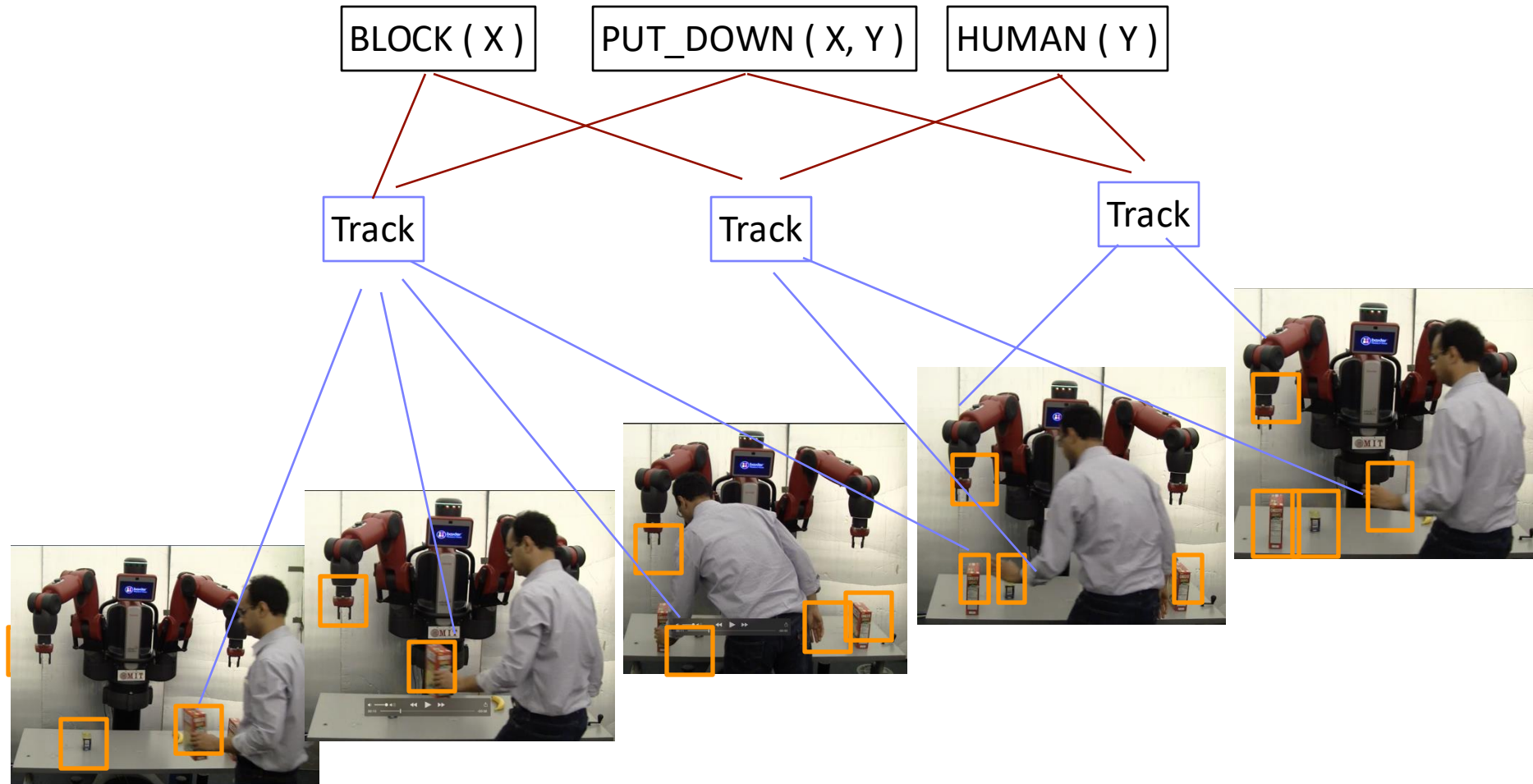
Sentence Tracker: Grounding Language in Vision

A. Barbu, et al. 2015

“the block that that the person put down”

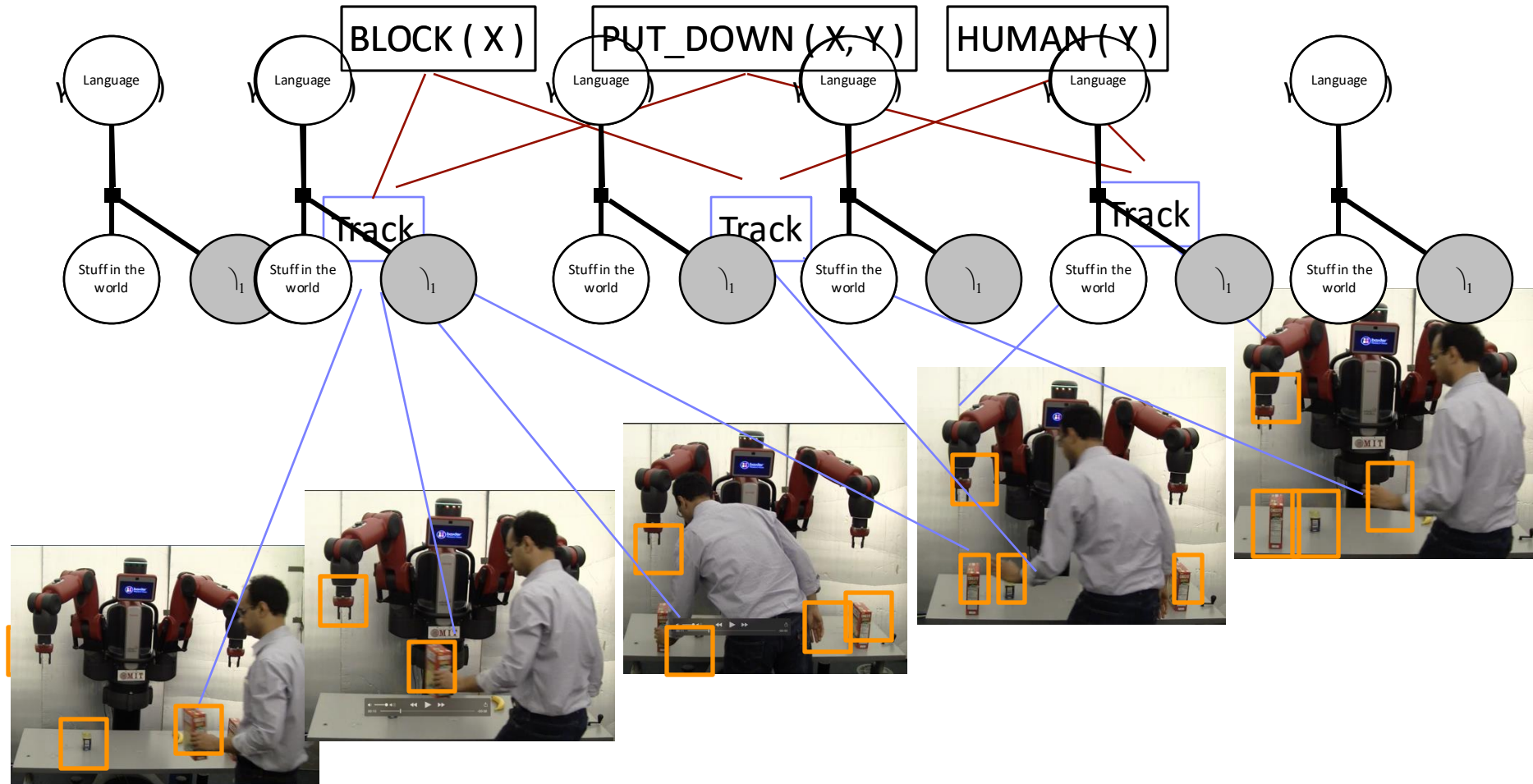


Solved as a factorial-HMM
Infers which object tracks fill the
roles in predicates

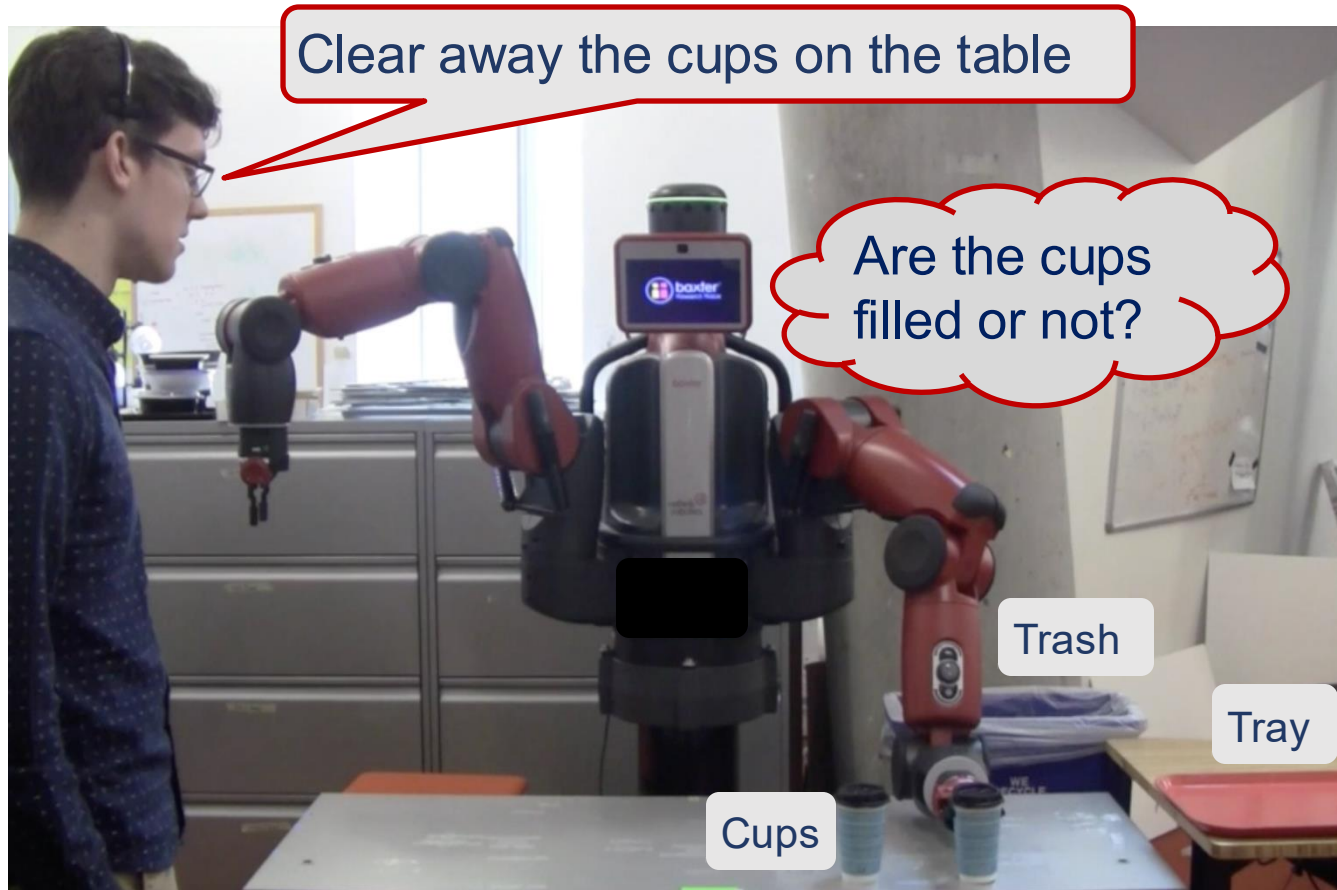


Re-interpretation as Grounding Graphs

“the block that that the person put down”



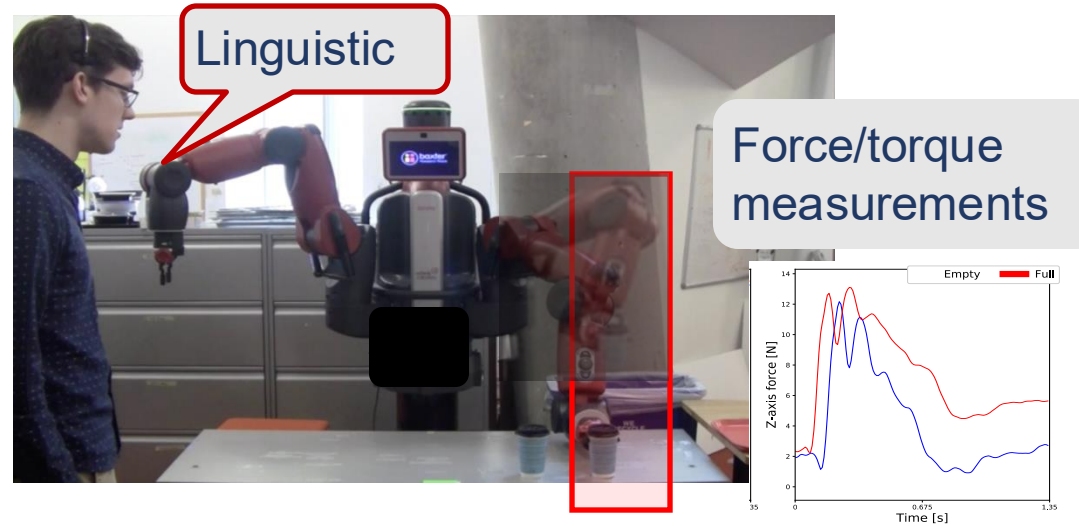
Need for Information Gathering



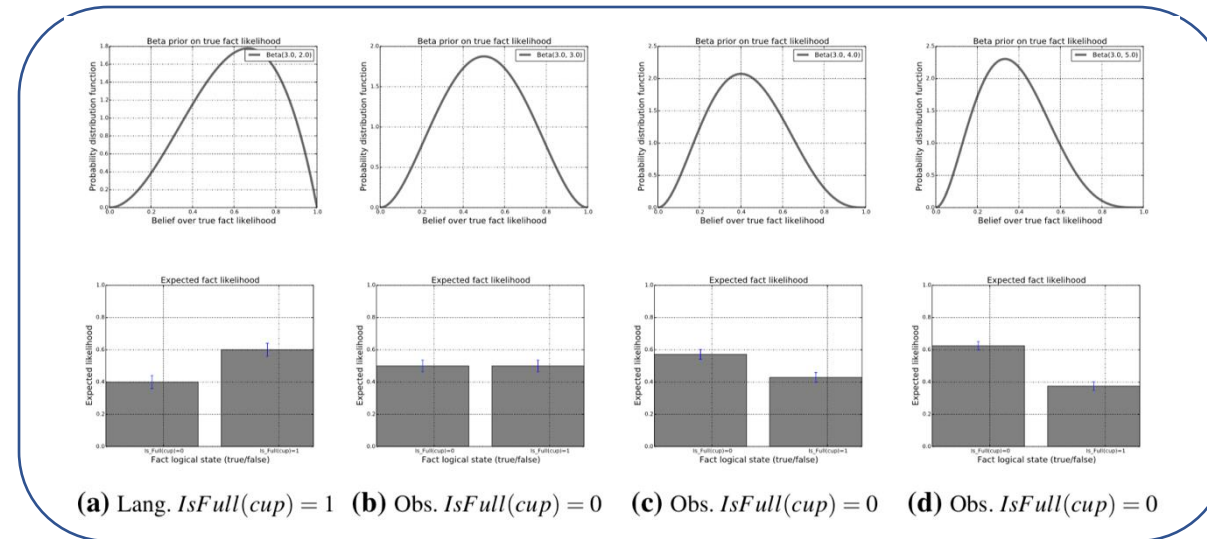
Missing knowledge causes failure in plan synthesis

Estimating Latent Symbolic States

Observations



Latent belief over object states



“Filtering” over discrete state variables using Beta-Bernoulli conjugate model.

Open Questions

- Long duration language understanding - understanding long missions
- Long-context multi-modal memory
- Dialogue systems – lack of (multi-modal) data

Open Questions

- NLP also includes speech. Robustness of speech understanding and working directly with speech.
- Natural human-like speech with emotion (contextually grounded)
- Explaining your actions
- New platforms – legged robots, humanoid robots.

Acknowledgements

Prof. Nicholas Roy
Prof. Thomas Howard
Prof. Daehyung Park

Prof. Parag Singla
Prof. Souvik Chakraborty
Dr. Subhro Roy

Prof. P. V. M. Rao
Prof. M. Balakrishnan
Dr. Piyush Chanana
Shri Dipendra Manocha

Namasivayam K, Gagandeep Singh, Rachit Thukral,
Himanshu Gaurav Singh, Vishal Bindal, Arnav Tuli,
Saptarshi Dasgupta, Akshat Gupta, Garima, Sandeep
Zachariah, Shubhangi Nema, Moksh Malhotra, Aman
Tambi, Nadir Mohd.

Ms. Sunita Negi
Mr. Shailendra Negi
Office teams: CSE, SIT, ScAI

Funding: DST, ICMR, MOD, IIT-IRD, MHRD



Thank you!

Your Feedback and Questions.