NRI: Hierarchical Representation Learning for Robot Assistants

Shuran Song, Carl Vondrick, Zhou Yu
Columbia University

Award ID: 2132519
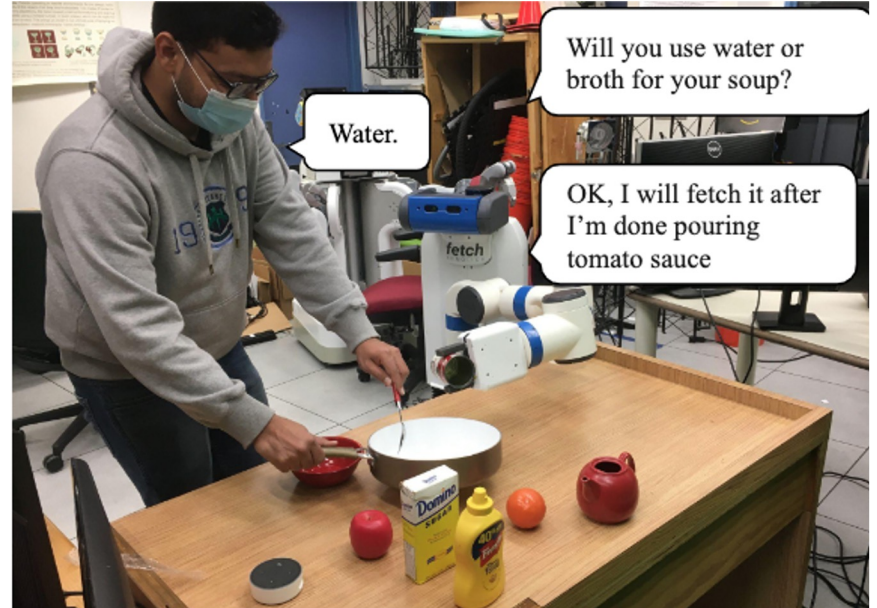
Project Goal:

Introduce **just-in-time robot assistant**.

From both **sight** and **dialogue**, our framework is able to anticipate what objects a person will need in the near future, and deliver it at exactly the right moment.

Applications for Just-in-time Object delivery

Kitchen
pass the proper ingredients to a chef

Repairment Shop
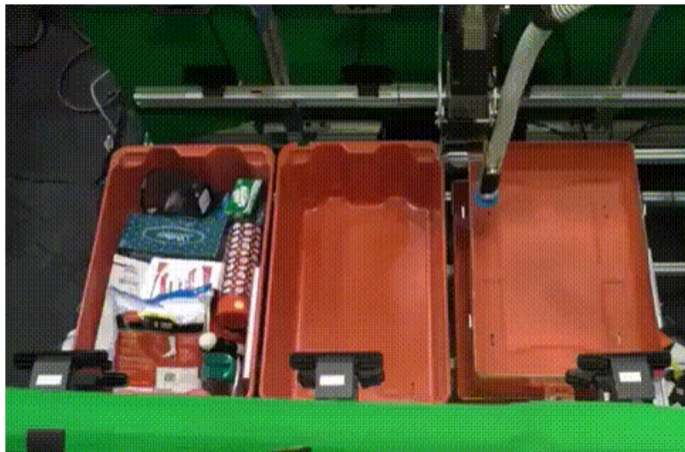pass the proper tool to a mechanic

Emergency room
pass the right equipment to a doctor

Challenges

What robots can do today

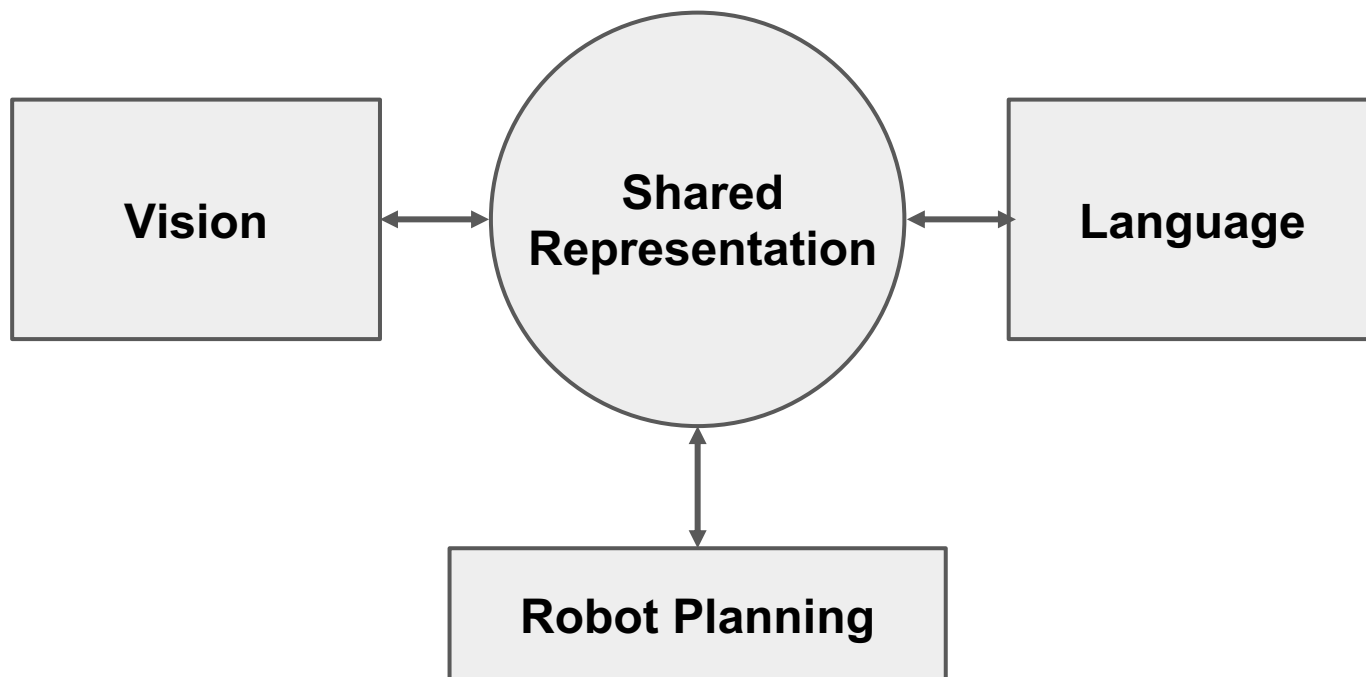Pick-and-Place a variety of objects

Navigate in indoor environments

Open challenge: Knowing **"what object"** to deliver at **"which time"** . Especially for tasks involve <u>long-horizon</u> planning and <u>ambiguous task procedures</u>.
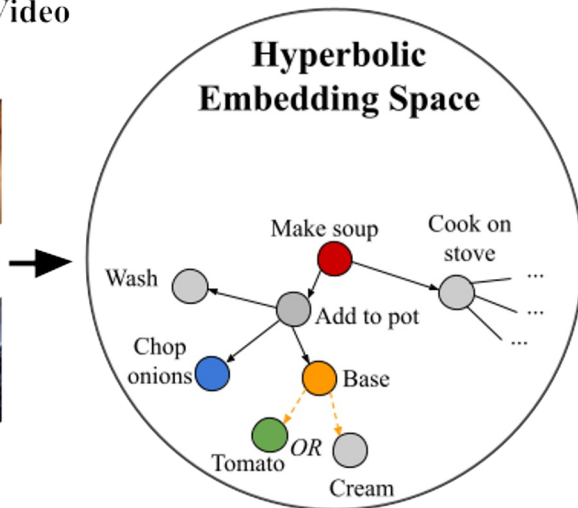
Approach:

Learning a Shared Representation for vision & language & robot planning
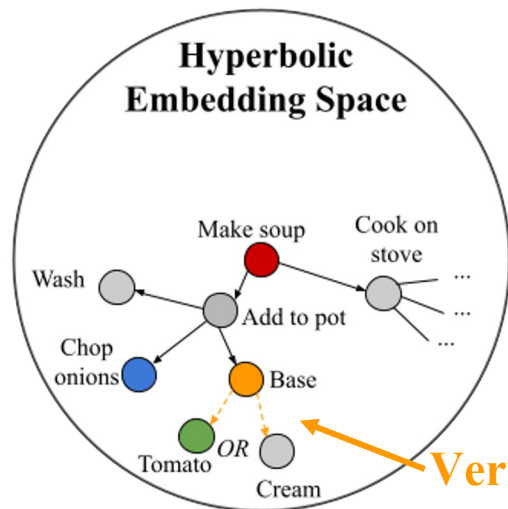
A: Learning Hierarchical Task Representations

Representation Learning from Video (Thrust A)

Hyperbolic Embedding Space

Encodes temporal hierarchy and dependencies between different states (i.e., goals or steps)

B: Understand User's Instructions and Intentions
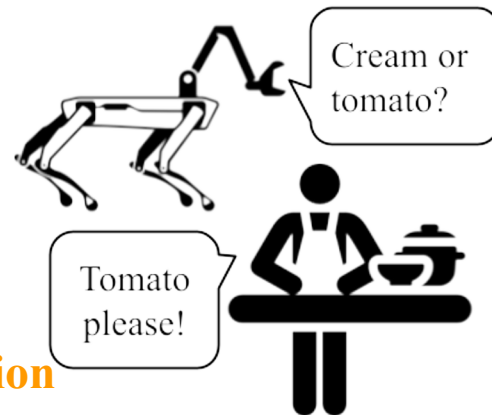
C: Intention-aware Planning
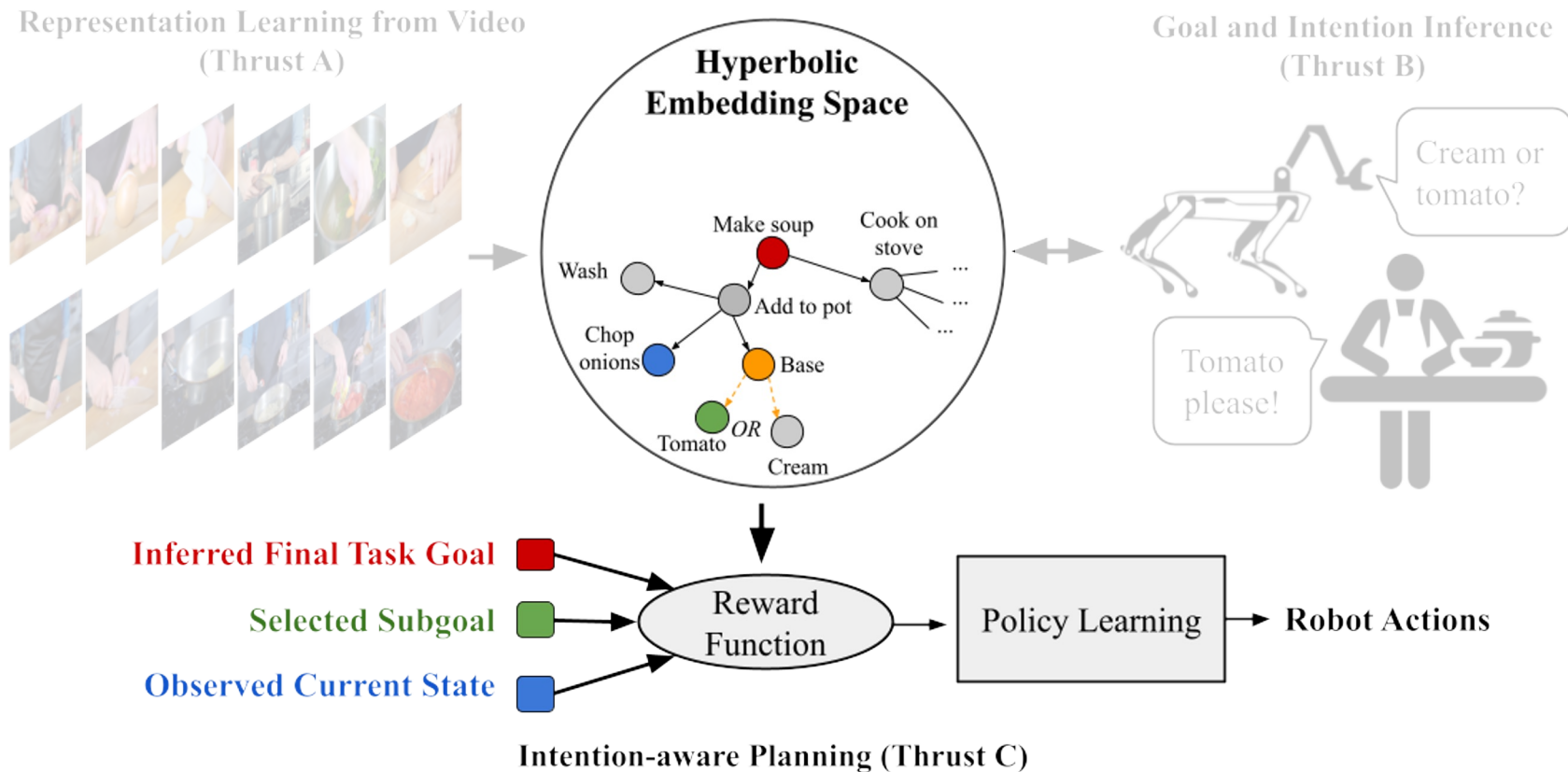
Border Impact

Expected results will enable machines to better understand and collaborate with people

- Improving the quality of life for people with limited mobility and other physical disabilities.
- Improving productivity for professionals in the workplace

New educational opportunities at the intersection of robotics, computer vision and nature language processing through a series of systematically designed curriculum and annual capstone projects for assistive robotics.