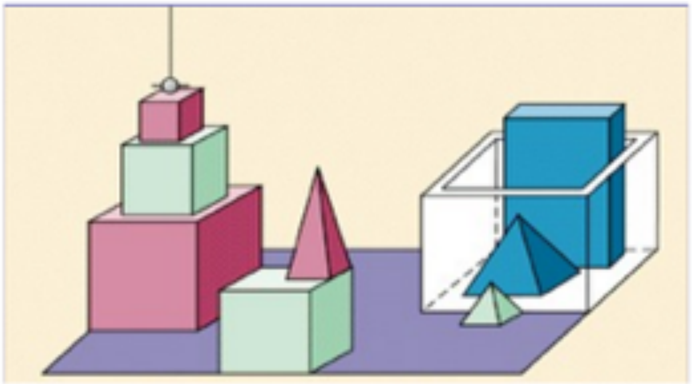
# ReGROUND:
Relational Symbol Grounding through Affordance Learning

# Winograd's SHRDLU [1968-70]

Micro-world in simulation:

- Dealing with blocks
- Natural language
- Simple planning and reasoning



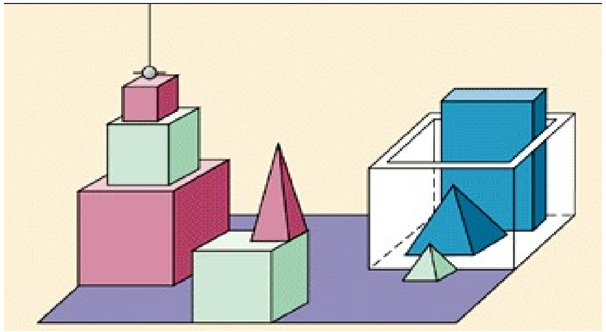
**Put the blue pyramid on the block in the box**

**Assumed: symbol grounding problem solved**

# Reground: Winograd's setting in the real world

**Bring me the tea pot and the sugar**

# From Winograd's SHRDLU to the real world



**Put the blue pyramid on the block in the box**

**Bring me the tea pot and the sugar**

- Three assumptions:
    - multi-modal input (perception and language); anchoring and symbol grounding
    - take into account the context & environment — multiple objects (multiple symbols & referents) and their relationships (relational representations)
    - build on a notion of affordance from robotics (captures relevant relationships)

KOÇ UNIVERSITY

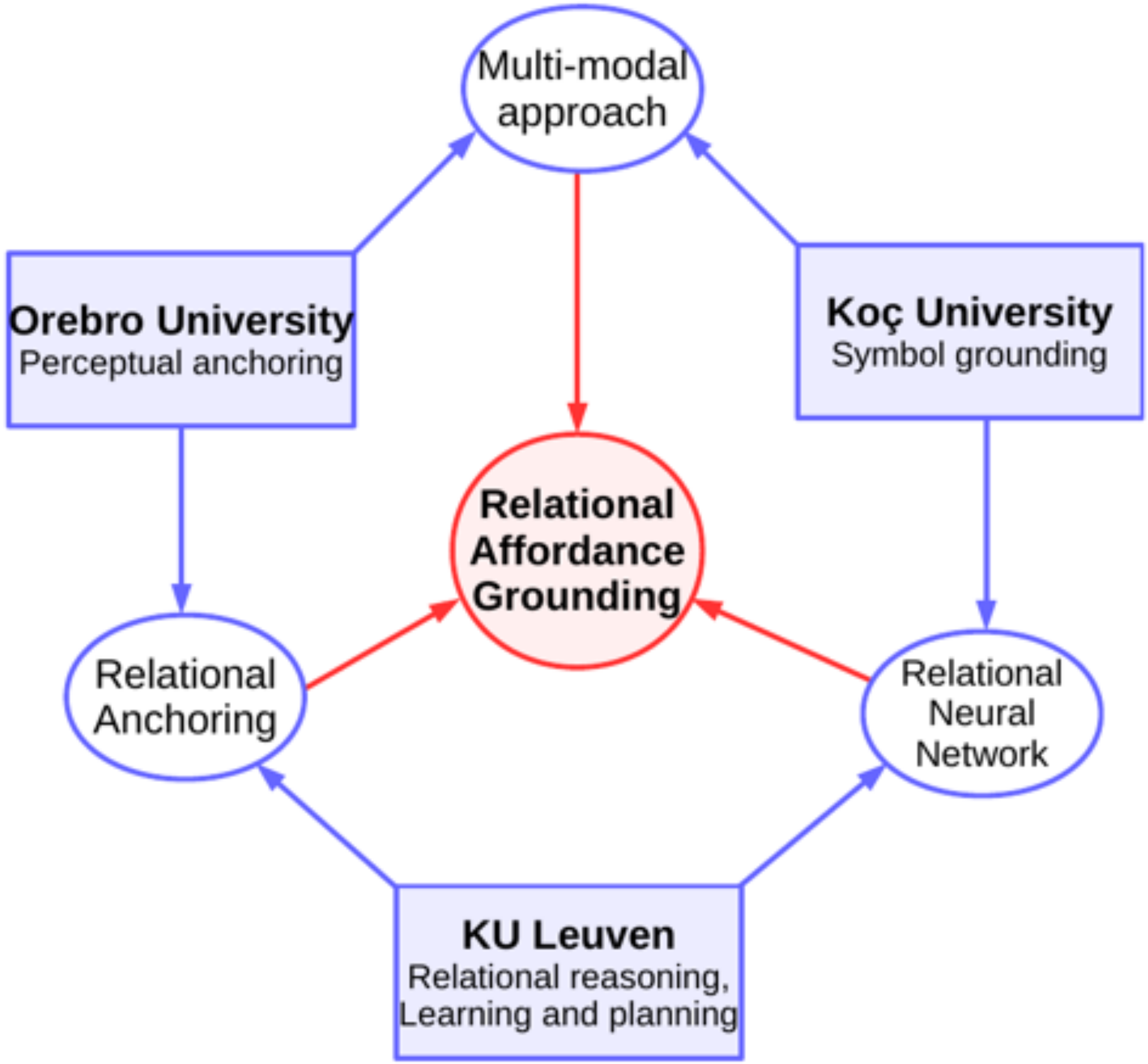
KU LEUVEN

ÖREBRO UNIVERSITET

# ReGROUND: experimental setup

Real World SHRDLU:
pick the cup right of the red apple

# The ReGROUND symbiosis



# Koç University

- Private research university with English instruction.
- Funded 1993 in Istanbul
- 4000 undergrad, 600 grad, 400 faculty
- Ranked $36^{th}$ among young universities by Times Higher Education.
- Among top 3 Turkish universities in publication rate and grants awarded

# Koç University ReGROUND Team

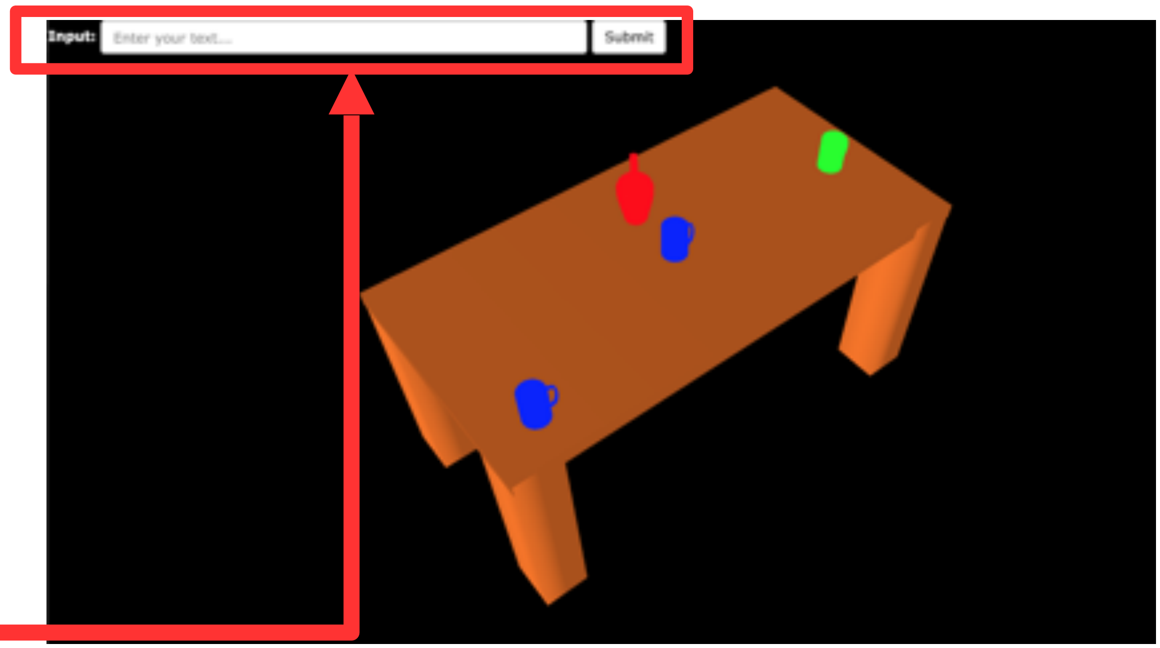
- **Deniz Yuret, Ozan Can, Emre Unal:**
  NLP, ML

- **Yücel Yemez:**
  Vision, Graphics

- **Aylin Kütay & Tilbe Göksun:**
  Language development

# **Koç University**
# Simulator for Language Grounding

- Mimic kitchen environment (color, size, position, type)
- Find object that is targeted with language in a 3DGrid world
- Text box for user to enter their instructions: *"Pick the blue mug next to the red bottle"*



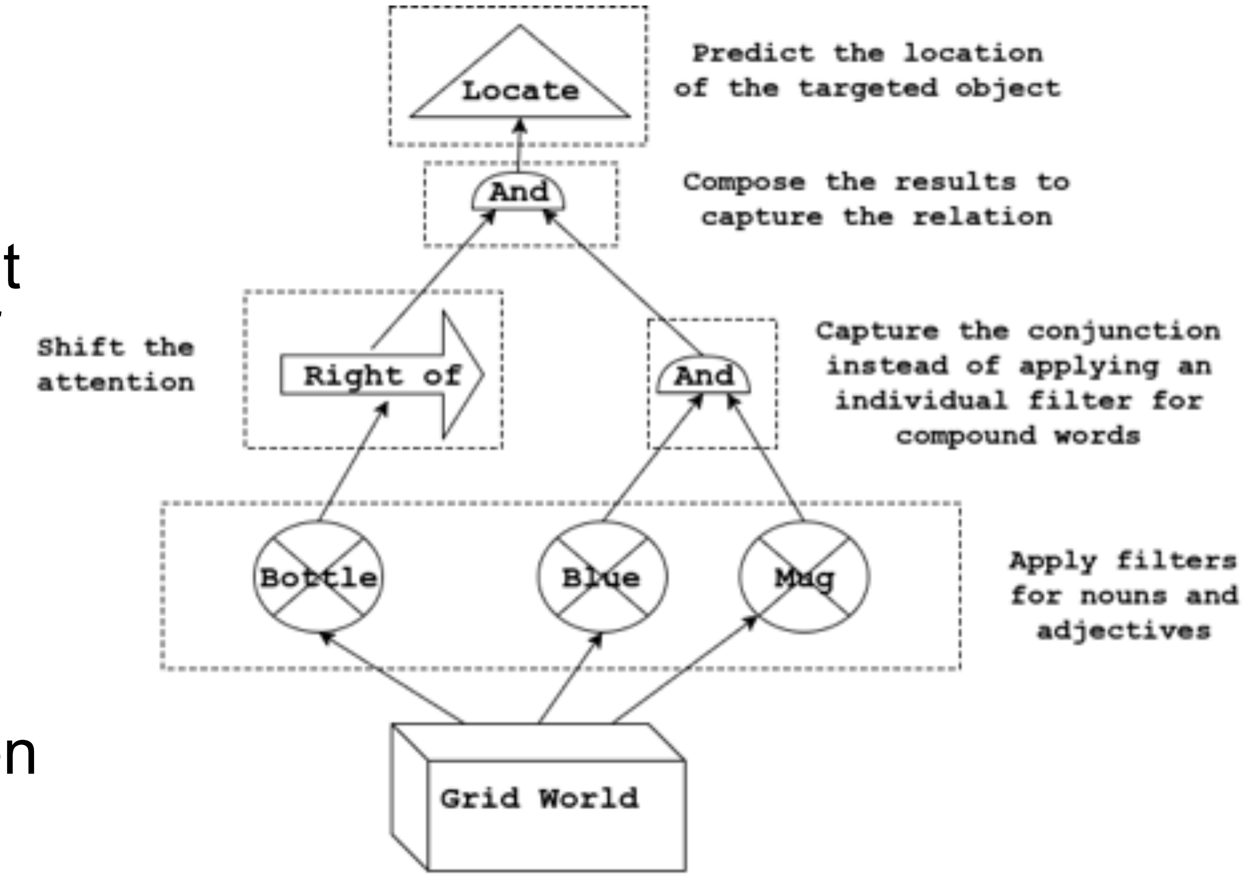

- **Goal:** learning to understand spatial instructions

# Koç University
## Symbol Grounding in Grid World

How to localize objects:

1. Generate network layout based on parsed tree of natural language instructions

2. Build neural network based
   on the components given by layout and predict position
   of targeted object




The layout generated for the instruction:
Pick the blue mug on the right of the bottle.

# Örebro University

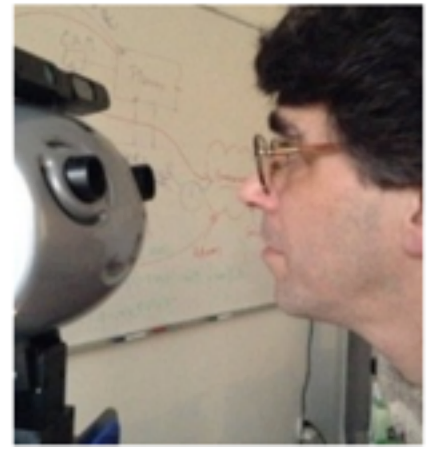
- University status since 17 years

- 17000 undergrad,
  500 doctoral students,
  1300 staff

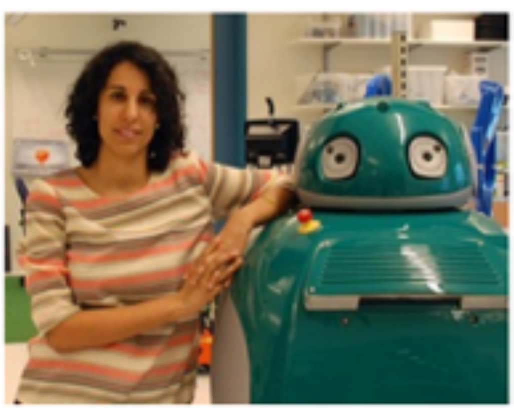
- Ranked 63rd among young universities by Times Higher Education.

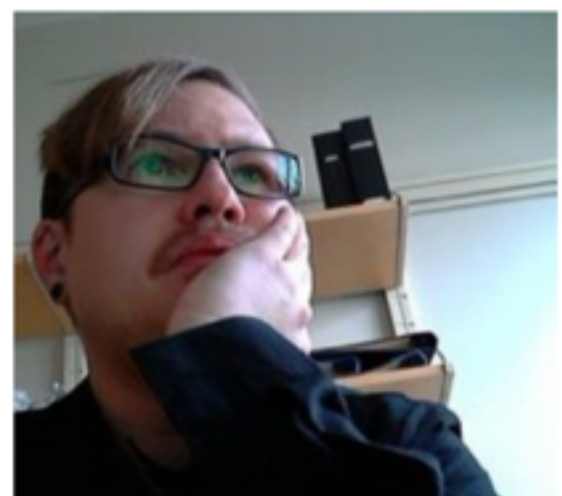
# Örebro University ReGROUND Team

- **Alessandro Saffiotti:** robotics, AI



- **Amy Loutfi & Andreas Persson:** anchoring



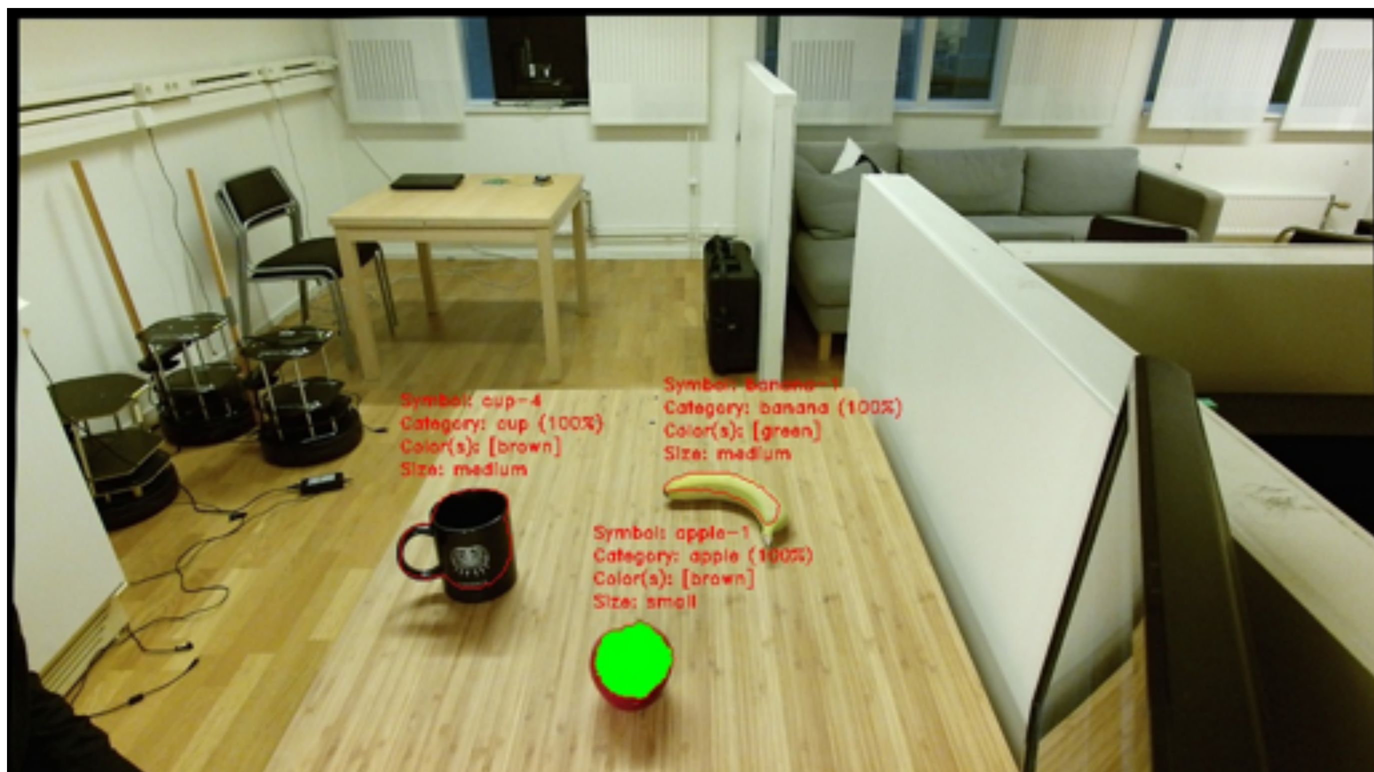


# **Örebro University**
# Perceptual Anchoring

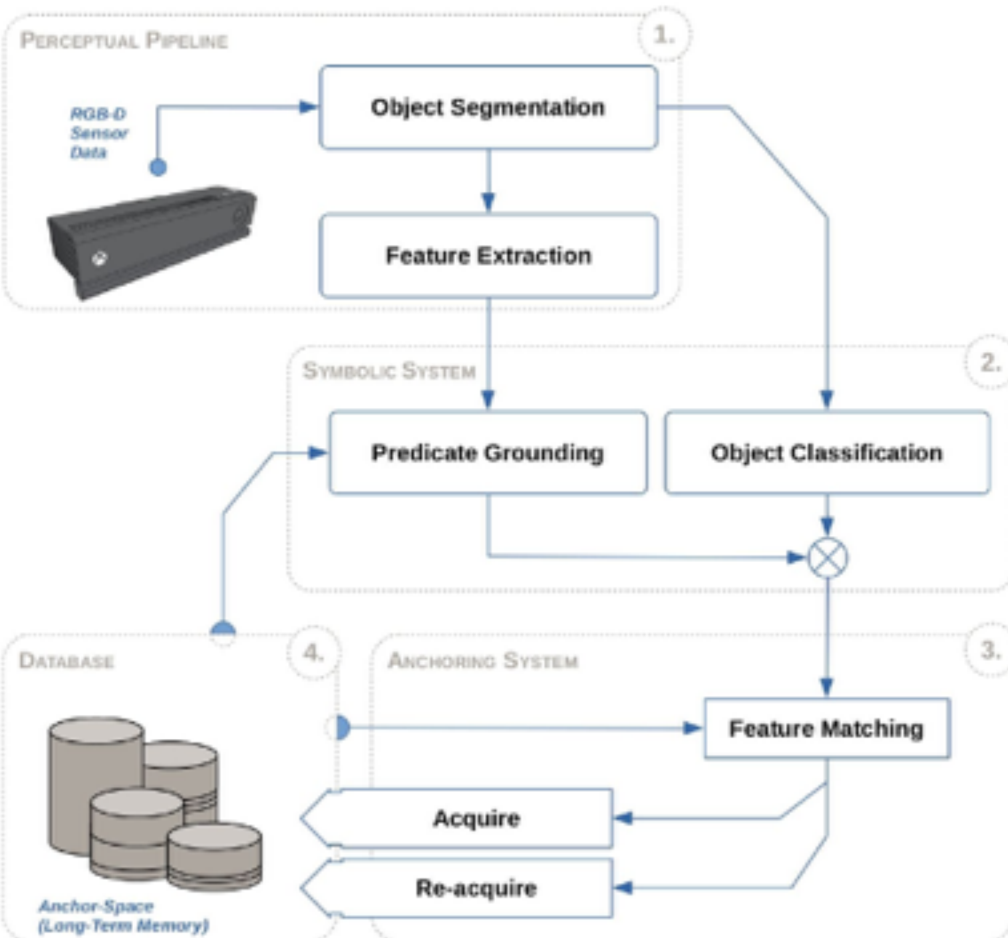
*“Anchoring is the problem of connecting, inside an artificial system, symbols and sensor data that refer to the same physical objects in the external world.”*

S. Coradeschi and A. Saffiotti. Robotics and Autonomous Systems 43(2-3):85-96, 2003.

Relate the symbol “**cup-4**” to the percept of the black mug over time



# Örebro University
# Perceptual Anchoring




1. **Perceptual Pipeline:**
   processing sensor data,
   segment objects,
   detect features
2. **Symbolic System:**
   establish the percept-symbol correspondence for segmented objects
3. **Anchoring System:**
   create and maintain anchors
4. **Database:**
   facilitates long-term maintenance of anchors

# KU Leuven

- Founded in 1425
- 50 000 student,
  10 000 staff
- Almost all disciplines represented

# KU Leuven ReGROUND Team

- **Jesse Davis:** ML, NLP
- **Luc De Raedt:** AI
- **Davide Nitti, Laura Antanas, Pedro Zuidberg:** robotics

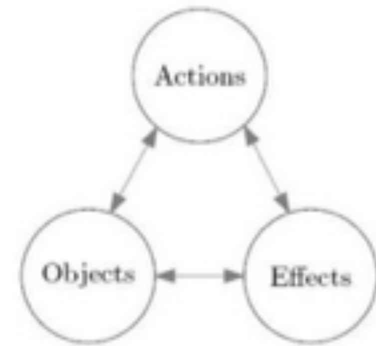




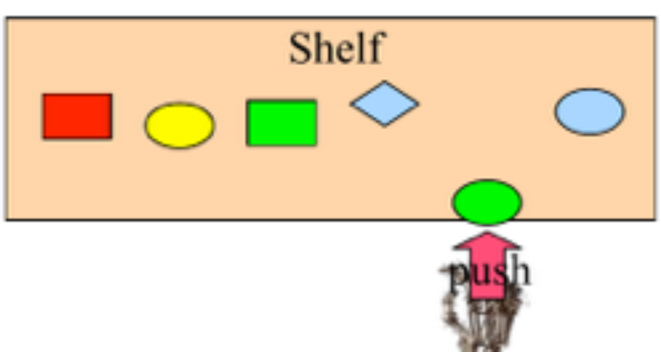
# KU Leuven
# Relational Affordances

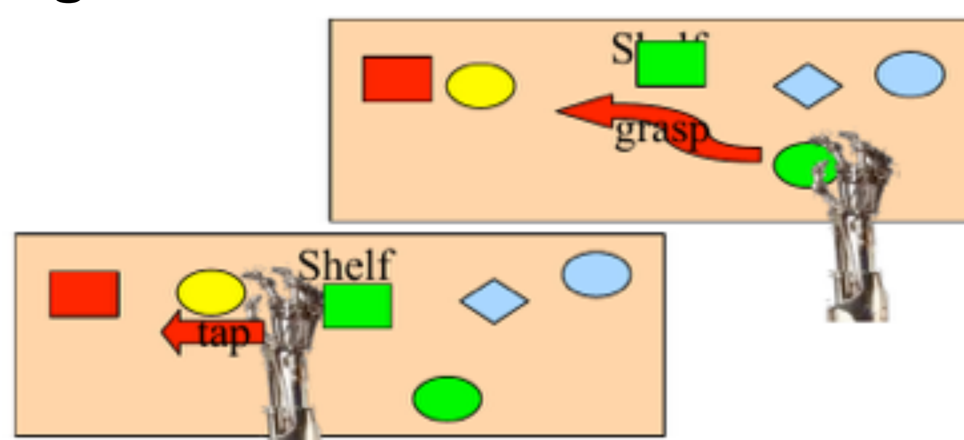
- **Object Affordance:**
  What can one do with particular object?
- **Relational Affordance:**
  in a particular context?

  with multiple objects and relations among them
- Use of statistical relational learning, **probabilistic programming for learning, reasoning and planning !**



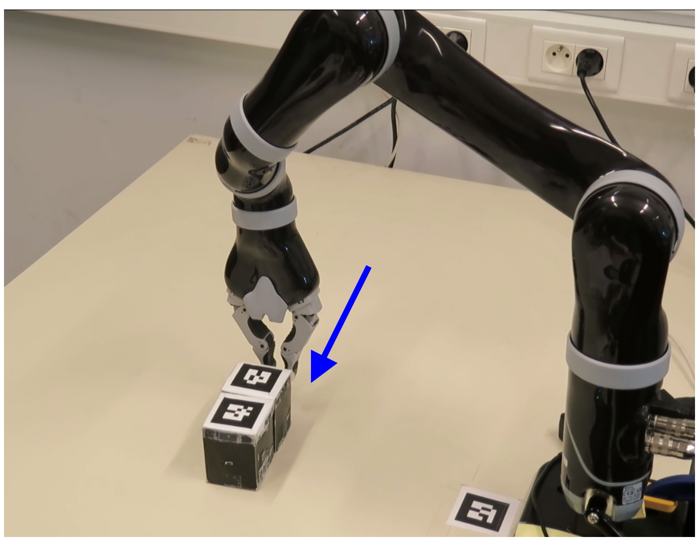

| Inputs | Outputs | Function |
|---|---|---|
| $(O, A)$ | $E$ | Effect prediction |
| $(O, E)$ | $A$ | Action recognition/planning |
| $(A, E)$ | $O$ | Object recognition/selection |



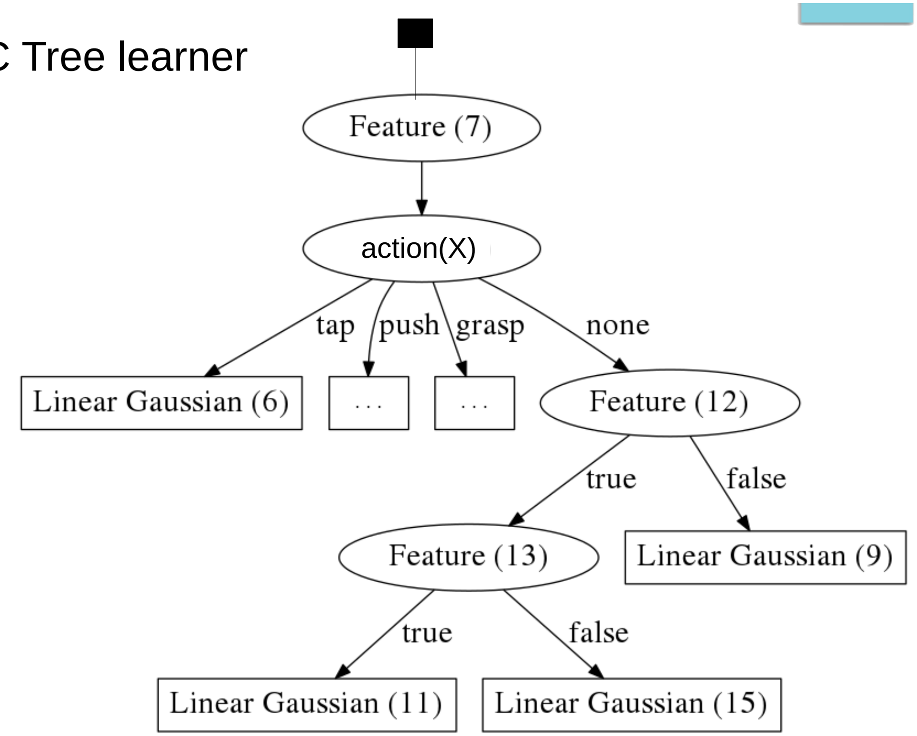



# KU Leuven
# Relational Affordance Learning

- **Learning the Structure of Dynamic Hybrid Relational Models**
  Nitti, Ravkic, et al. ECAI 2016
  - Captures relations/affordances
  - Suited to learn affordances in robotics set-up, continuous and discrete variables
  - Planning in hybrid robotics domain

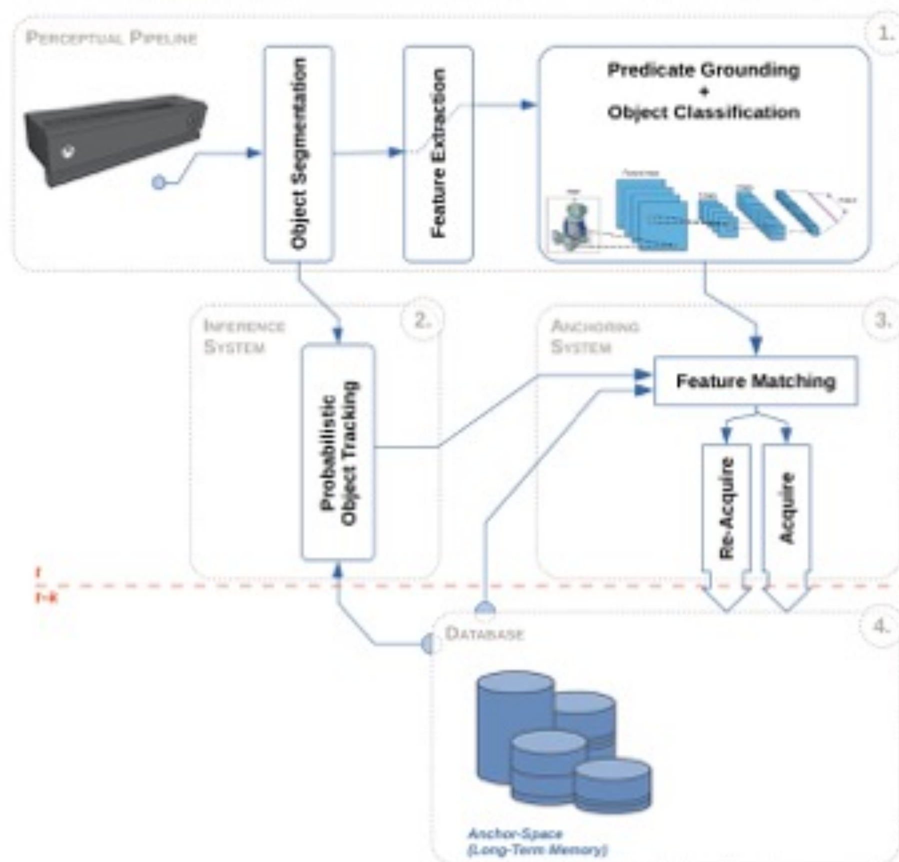


DDC Tree learner



# From Object Anchoring to Relational Anchoring

- Use relational object tracker
- Use inference system as short-term memory
- Use anchor database as long-term memory
- Maintain cognitive consonance between the two systems
- Leverage relational affordances to improve anchoring
  → **from object anchoring to relational anchoring**
  → **support planning and reasoning**



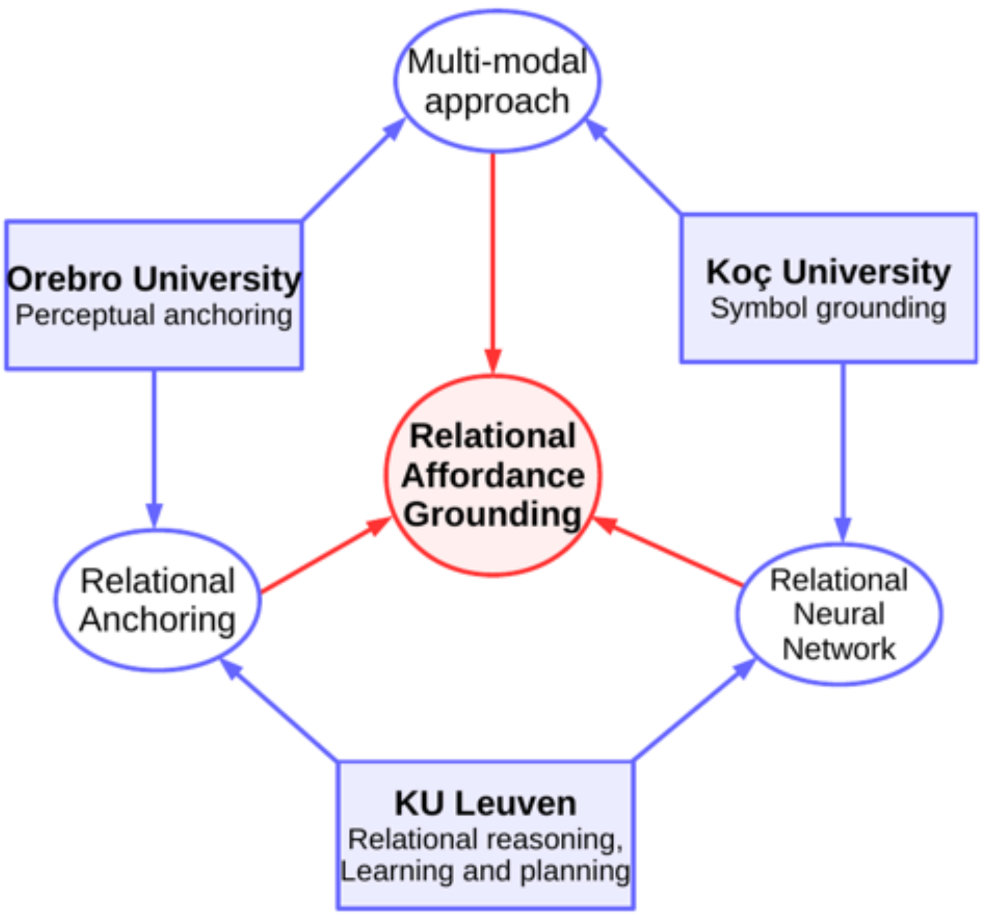
# Ongoing Work

- Replace the simple feed forward architecture for the neural network in the symbol grounder with an attention based sequential architecture
  → facilitate **grounding of verbs and prepositions**
- Compare human attention mechanism in eye movement to attention based neural nets
  →**improve architecture**
- Define/learn relational affordances to improve anchoring of objects

# Targeted Breakthrough

**From “symbol grounding” to “affordance grounding”**

agents can adapt to new environment and interpret possibly unimodal input to achieve goals

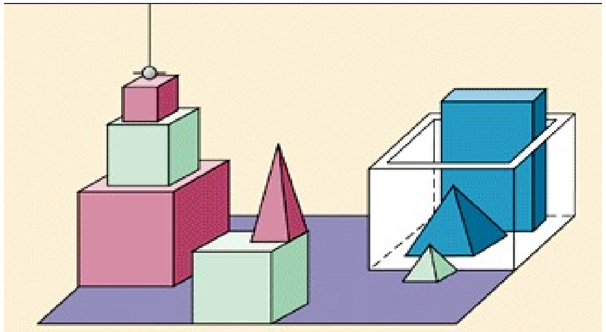


# ReGROUND: experimental setup

Real World SHRDLU:

pick the cup right of the red apple

# From Winograd's SHRDLU to the real world



**Put the blue pyramid on the block in the box**

**Bring me the tea pot and the sugar**

- Three assumptions:
  - multi-modal input (perception and language); anchoring and symbol grounding
  - take into account the context & environment — multiple objects (multiple symbols & referents) and their relationships (relational representations)
  - build on a notion of affordance from robotics (captures relevant relationships)

KOÇ UNIVERSITY

KU LEUVEN

ÖREBRO UNIVERSITET

# Expected Outputs

1. Novel techniques that perform symbol grounding by learning relational affordances.

2. New insights into symbol grounding, relational affordances, and developmental language learning.

3. Demonstration that relational grounding significantly improves the percentage of successfully completed tasks in both a simulated environment and a real robotic setup about a household environment.

4. Publicly available data from the robotic and simulated environment and code for our algorithms.