



**Center for
Integrated
Cognition**

Soar Learning to Make Breakfast

Robots Collaborating with Humans
Using Natural Language
in the
Real World

45th Soar Workshop

Peter Lindes

5 May 2025

Assisted by Kaoutar Skiker

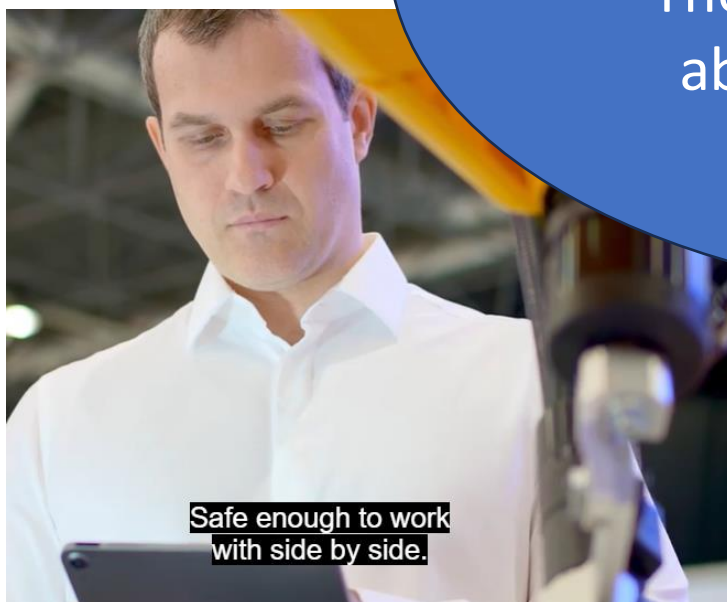


Robot Assistants

ated Robotic Therapy Dogs & Cats



- What's missing in all of these?
- The ability to interact using language natural to humans
 - The ability to learn and reason about a large, complex body of situational knowledge





Objective: Robot assistants with language



Imagine hiring a human cook and giving them a lot of training.

Then imagine replacing that person with a robot.

- Hard problems:

- Vision
- Manipulation
- Navigation
- Knowledge and reasoning
- Language understanding

- Existing approaches:

- Classic ITL
 - Limited language
 - Limited knowledge
- Agentic LLMs
 - Not grounded
 - Limited reasoning



Use Case: A Kitchen Assistant (a private cook)

Tasks:

- Prepare meals
- Serve meals
- Clean up
- Maintain inventory
- Make shopping list
- Order groceries
- Store groceries



Knowledge needed:

- Kitchen layout
- Storage locations
- Ingredients
- Recipes
- Tools
- Cleaning procedures
- Personal preferences
- Current inventory
- Ordering procedures

You: I'm hungry. What's for breakfast?

Cook: What would you like for breakfast today?

You: Well, maybe some toast and an omelet.

Cook: You like your toast buttered and with jam, right?

You: Sure, sounds good.

Cook: We're out of cheddar cheese. Is Swiss OK?

You: Well, I guess that'll do. Be sure to get some cheddar!

Cook: Will do.

(Goes off and prepares the breakfast and serves it.)

Cook: Here's your breakfast. **You:** Great, thanks!

(Cleans up the kitchen and orders cheddar cheese.)



A New

Agent:
Integration

Knowledge:
Situation
Preferences
Tasks

Human:
Overall direction
Knowledge from
interaction

Robot:
Perception
Action

LLM:
Language
translation
General
knowledge

We have
experience in all
these areas.





Three Dimensions of the Problem

- Making free-form language understandable to the agent
- Using interactions to learn the knowledge for complex tasks
- Generalizing this knowledge over multiple tasks



Request to ChatGPT: (2/20/2025, 11:18am MST)

I need a recipe for making scrambled eggs. Please give it to me as a list of a few simple steps.

Response from ChatGPT:

Here's a quick and easy recipe for scrambled eggs:

Ingredients:

- 2–3 eggs
- Salt (to taste)
- Pepper (to taste)
- 1 tbsp butter (or oil)
- Optional: Milk or cream (1–2 tbsp for creamier eggs)

Steps:

1. **Crack the eggs** into a bowl and beat them with a fork or whisk until fully combined.
2. **Add a pinch of salt** and pepper to the eggs. Optionally, add a splash of milk or cream for a creamier texture.
3. **Heat a non-stick pan** over medium-low heat and add the butter.
4. **Pour the eggs into the pan** once the butter has melted, swirling the pan to coat the bottom.
5. **Cook gently**, stirring occasionally with a spatula, until the eggs are soft and slightly runny or fully cooked to your liking.
6. **Serve immediately** and enjoy your scrambled eggs!

Let me know if you'd like to adjust the recipe to your taste!

Example of Knowledge from LLM

Agent must learn:

- Ingredients
- Measures
- Actions
- Manners
- Criteria

Translation to formal representations

Human interaction for details and verification



Issues in Grounded Language Understanding

- Non-standard language
 - Free form, ungrammatical
 - Indirect
 - Abbreviated
- Many alternatives
- Object references
- Verb meanings
- Abstract concepts
- Individual variations



Challenging References

Spatial relations

Discourse properties

Goal: Put a **cooked** potato slice on the counter.

Goal: Place a slice of **cooked** potato onto the counter.

Goal: Put a piece of **cooked** potato on the counter.

The knife

yellow handled knife

The knife **from in between the lettuce and the apple**

The knife

the knife **used to cut the potato**

A slice in the potato

Potato slice

The potato slice

The potato

A slice of potato

The **cooked** potato slice

The potato **on the gray table**

A slice of **the cut potato**

The potato slice

round table **with tablecloth**

The table

The gray table

counter

counter **with blue plate on it**

the counter **left of the bread**

the counter **near the soap container**

the counter **in front of the stove**

the counter left of the bread

the counter

the kitchen counter **with the loaf of bread**

the counter top with the loaf of bread

Examples from ALFRED dataset:

Shridhar et al. (2020), *ALFRED: A Benchmark for Interpreting Grounded Instructions for Everyday Tasks*. arXiv:1912.01734v2



Learn to “cook”

Cook <something>.

- Cook a potato.
- Cook a potato in the microwave.
- Make my breakfast.
- Prepare dinner for the family.
- Fix me an omelet.
- Cook eggs benedict.
- Boil an egg (*hard or soft?*)

Issues to consider

- Overall context
 - Who will be eating, preferences
 - What, where and when
 - Presentation and serving
 - Cleanup and maintenance
- Results: entities ready to eat:
 - An individual food item (*toast*)
 - A dish (*scrambled eggs*)
 - A meal (*breakfast*)
- Verbs in context
- Learning



Making Breakfast Dialog

A dialog with a human cook

You: I'm hungry. What's for breakfast?

Cook: What would you like for breakfast today?

You: Well, maybe some toast and an omelet.

Cook: You like your toast buttered and with jam, right?

You: Sure, sounds good.

Cook: We're out of cheddar cheese. Is Swiss OK?

You: Well, I guess that'll do. Be sure to get some cheddar!

Cook: Will do.

(Goes off and prepares the breakfast and serves it.)

Cook: Here's your breakfast. **You:** Great, thanks!

(Cleans up the kitchen and orders cheddar cheese.)

Free-form
language
Full situational
knowledge

A dialog with a well-trained Thor-Soar

You: Make my breakfast.

Thor-Soar: Would you like your regular breakfast today?

You: Yes.

Thor-Soar: I will make and serve your breakfast.
(Goes off and prepares the breakfast and serves it.)

Thor-Soar : Your breakfast is served.

You: Thank you!

(Thor-Soar puts cooking tools in the sink.)

(You eat your breakfast.)

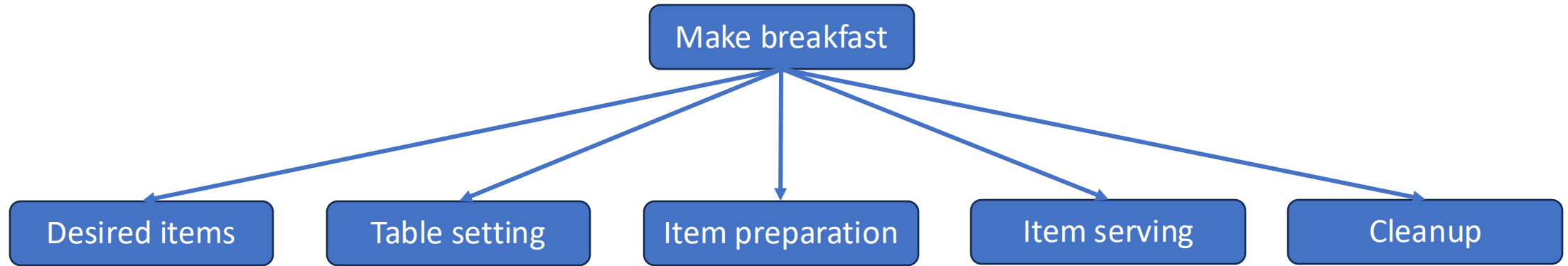
You: I am done.

(Thor-Soar cleans up the table, puts all the dirty dishes in the dishwasher and turns it on. When the dishwasher is done it puts all the dishes away in their places.)

Constrained
language
Learned situational
knowledge



Subtasks of Make Breakfast



Issues:

- Learning dialogs
- Knowledge representation
- Incremental generalization
- Interleaving subtasks



Learning *open*

I: Open the fridge.

TS: What are the preconditions for this action?

I: The fridge is reachable and you are not holding anything.

TS: What will this action accomplish?

I: The fridge is open.

TS: What is the name of the output command?

I: open

TS: Thank you, I will try to perform the action.



Examples of Generalization Issues

I: Open the fridge.

TS: What are the preconditions for this action?

I: The fridge is reachable and you are not holding anything.

What about
opening a
microwave that is
turned on?

I: Discard the apple.

TS: What is the goal of discard?

I: The goal is that the apple is in the garbage can.

Where should we
put an empty soda
can?

- Multiple action models or variations of a single one?
- Can Soar “forget” a chunk after it has been learned?



Nuggets

- Soar can manage complex tasks
- Soar can learn situational knowledge
- Soar can get an LLM to help it understand human language
- Soar can get general knowledge from an LLM
- Soar and an LLM can cooperate to make robotic assistants



- Using an LLM to make language understandable is a challenge
- Coordinating ITL and an LLM to learn situational knowledge is a challenge
- Grounding language to the physical world is a challenge