Large Language Models for Autonomous Robot Planning and Acting

Faculty candidate presentation at Rose-Hulman, ECE

Research talk (20 minutes) & Vision talk (20 minutes)

Yue Cao

Purdue University, ECE

yuecao@purdue.edu

Overview

Part 1. Research Talk

1. Real Autonomous Robots
2. LLM for Planning
3. LLM for Acting

Part 2. Vision Talk

1. AI for ECE
2. Teaching and Research

Fake Autonomous Robots

Kiwibot 2018: food delivery at UC Berkeley campus

What is actually behind?

San Francisco Chronicle

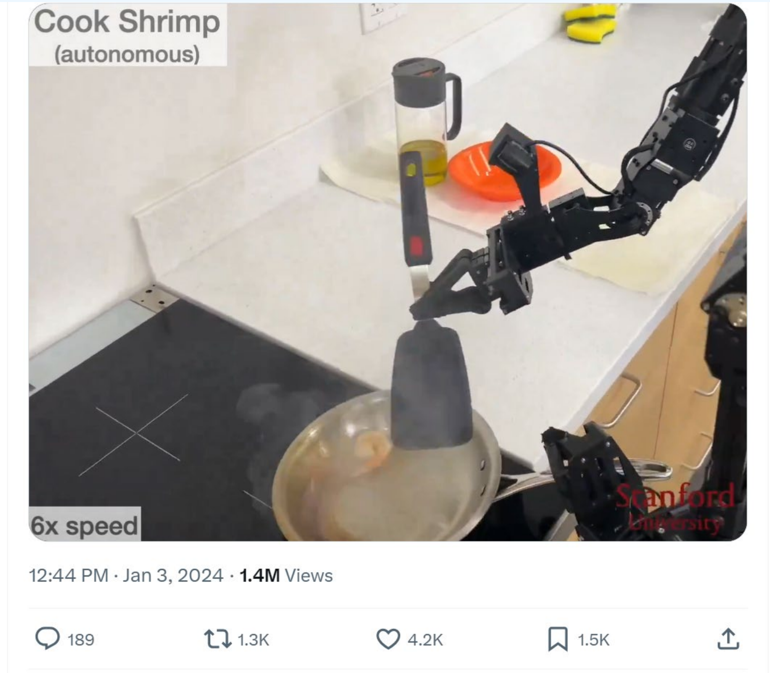
The Kiwibots do not figure out their own routes. Instead, people in Colombia, the home country of Chavez and his two co-founders, plot “waypoints” for the bots to follow, sending them instructions every five to 10 seconds on where to go.

As with other offshoring arrangements, the labor savings are huge. The Colombia workers, who can each handle up to three robots, make less than $2 an hour, which is above the local minimum wage.

“Kiwibots win fans at UC Berkeley as they deliver fast food at slow speeds”, San Francisco Chronicle, 2019

Fake Autonomous Robots

Mobile ALOHA 2024: housekeeping, Stanford Lab



What is actually behind?

IEEE Spectrum | Robotics

What's wrong with teleoperation, then?

Nothing! Teleoperation is great. But when people see a robot doing something and it *looks* autonomous but it's *actually* teleoperated, that's a problem, because it's a misrepresentation of the state of the technology. Not only do people end up with the wrong idea of how your robot functions and what it's really capable of, it also means that whenever those people see *other* robots doing similar tasks autonomously, their

"That Awesome Robot Demo Could Have a Human in the Loop", IEEE Spectrum , 2024

Towards Real Autonomous Robots

Two fundamental questions

1. New technology?
   Artificial intelligence (AI)

2. Robot capability?

Sense
↓
Plan
↓
Act

IEEE Spectrum | Robotics

What's wrong with teleoperation, then?

Nothing! Teleoperation is great. But when people see a robot doing something and it *looks* autonomous but it's *actually* teleoperated, that's a problem, because it's a misrepresentation of the state of the technology. Not only do people end up with the wrong idea of how your robot functions and what it's really capable of, it also means that whenever those people see *other* robots doing similar tasks autonomously, their

Towards Real Autonomous Robots

Two fundamental questions

1. New technology?
   Artificial intelligence (AI) →

2. Robot capability?

Sense → Plan → Act

My Research Focus

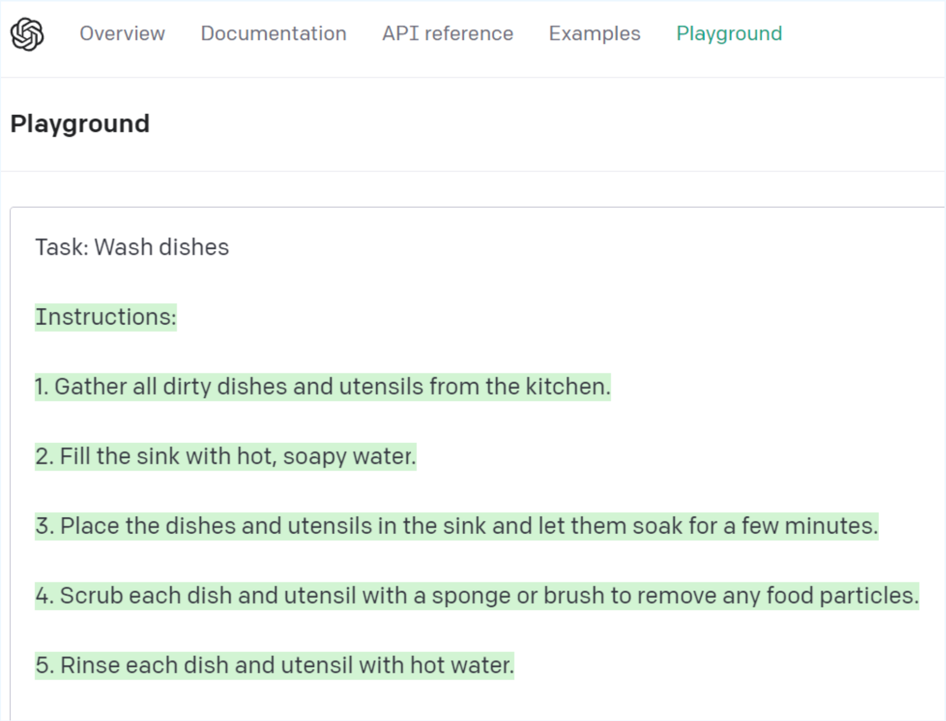
Large Language Models

For

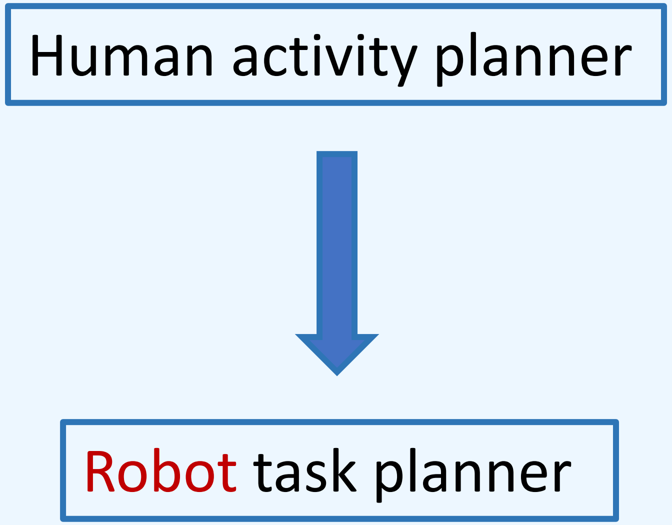
Planning and Acting

LLM for Planning

GPT can do planning for human activities



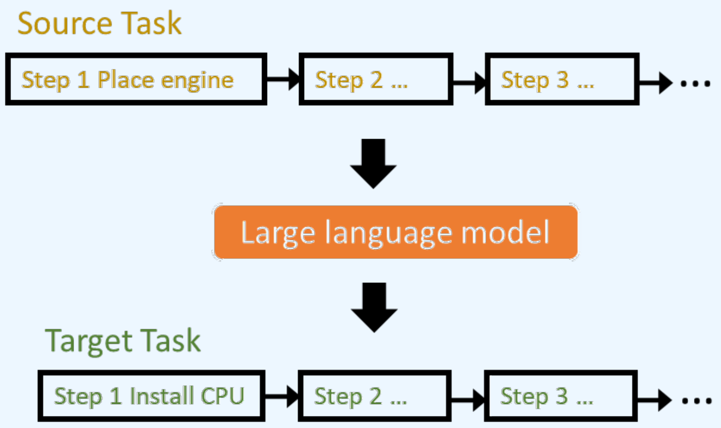
How about using LLMs in robot tasks?

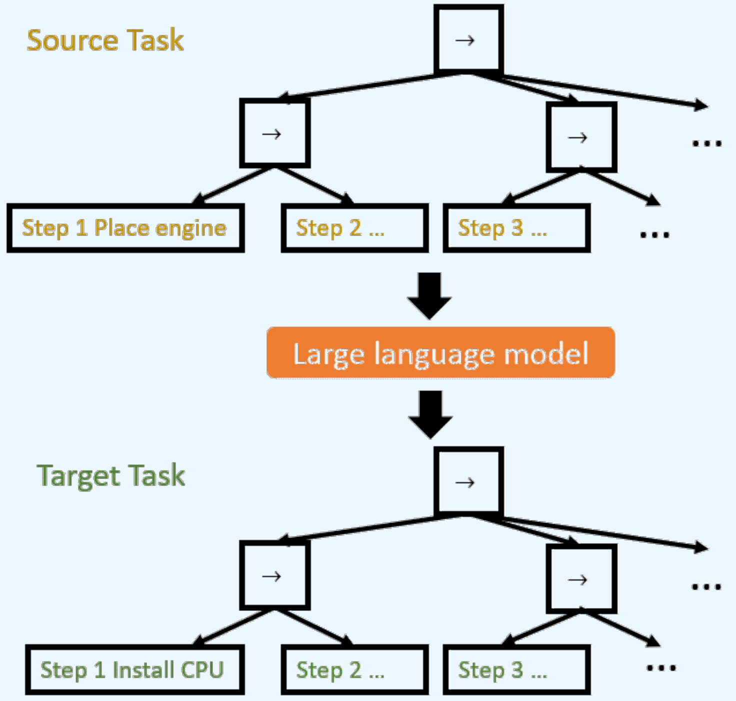


LLM for Planning

Previous works:
Generated task are sequential

Can we generate modular tasks?
Better reusability and readability

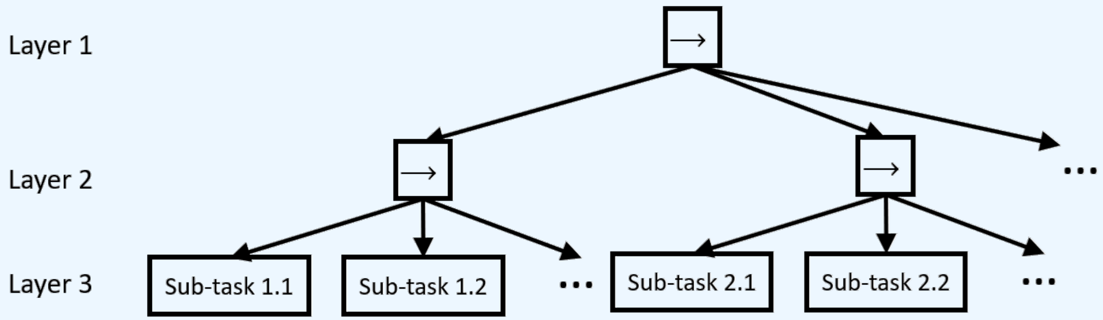




Yue Cao and C. S. George Lee, “Robot Behavior-Tree-Based Task Generation with Large Language Models”, AAAI Spring Symposium, 2023

LLM for Planning

- Behavior-Tree Task Generation with LLMs
  Prompt Design: Phase-Step Prompt



Source Task
Procedures:
Phase 1.
Step 1. [Sub-task 1.1] $_X$; Step 2. [Sub-task 1.2] $_X$; ...
Phase 2.
Step 1. [Sub-task 2.1] $_X$; Step 2. [Sub-task 2.2] $_X$; ...
...
Target Task: Task Description
Procedures:
[Phase 1.
Step 1. Sub-task ?; Step 2. Sub-task ?; ...
Phase 2.
Step 1. Sub-task ?; Step 2. Sub-task ?; ...
...] $_Y$

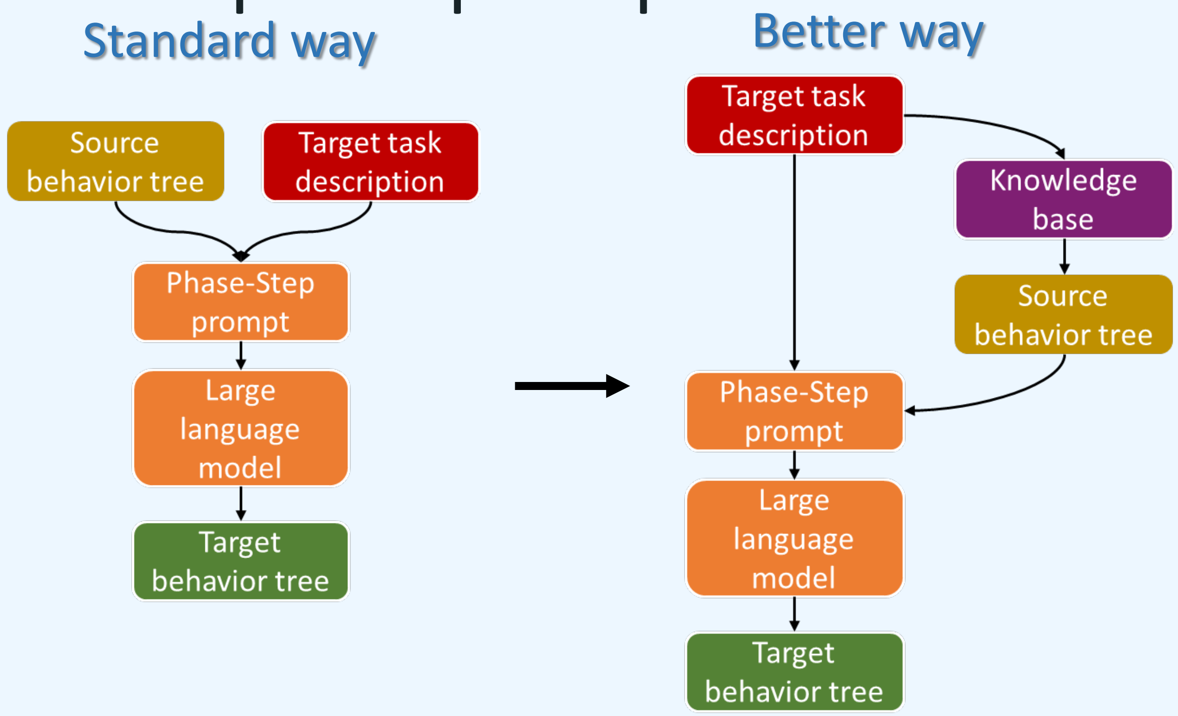
Layer 2 node: Phase; Layer 3 node: Step

LLM for Planning

- Use domain knowledge to improve prompt

Source Task
Procedures:
Phase 1.
Step 1. Put car at a conveyor;
Step 2. Lift the car.
Phase 2.
Step 1. Pick the wheel;
Step 2. Approach conveyor;
Step 3. Align wheel with wheel hub.
Phase 3.
Step 1. Insert screws;
Step 2. Fasten screws;
Step 3. Leave the conveyor.

Target Task: Desktop assembly
Procedures:
To be generated …



Integrate knowledge base retrieval into the pipeline

The end-user just needs to input a short target task description. The source behavior tree can be automatically retrieved from a knowledge base.

LLM for Planning

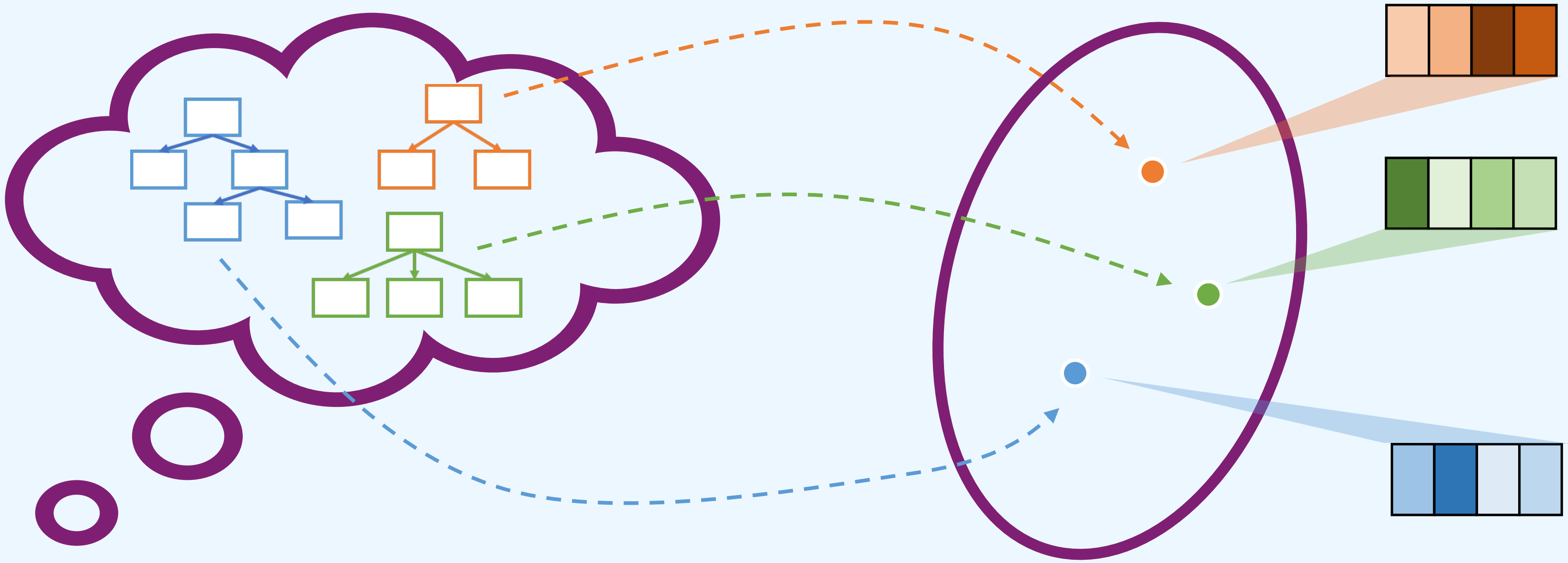
- How to retrieve similar tasks from a knowledge base?

- Solution:
  Embedding, mapping an entity into a fixed length vector.

word, document, Amazon product, Netflix movie…(Entity2Vec)

LLM for Planning

- Vector Embeddings for behavior-tree tasks



Knowledge base storing tasks

Vector space

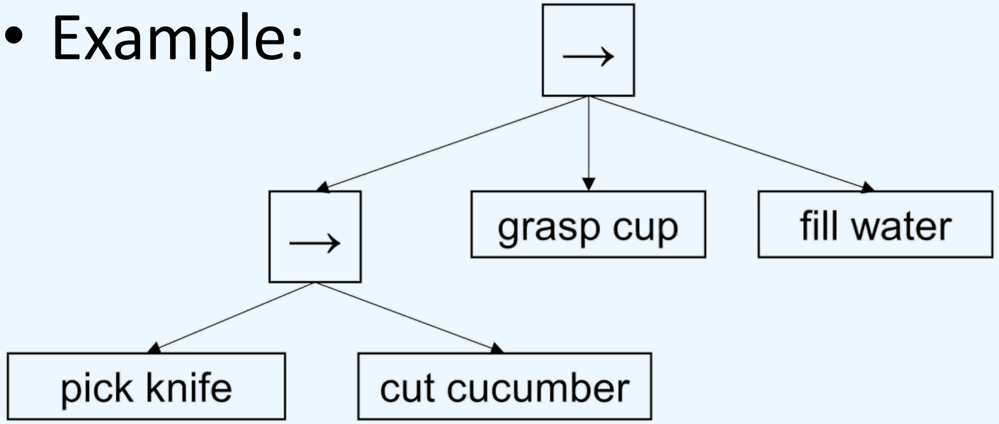
Yue Cao and C. S. George Lee, “Behavior-Tree Embeddings for Robot Task-Level Knowledge”, IROS, 2022

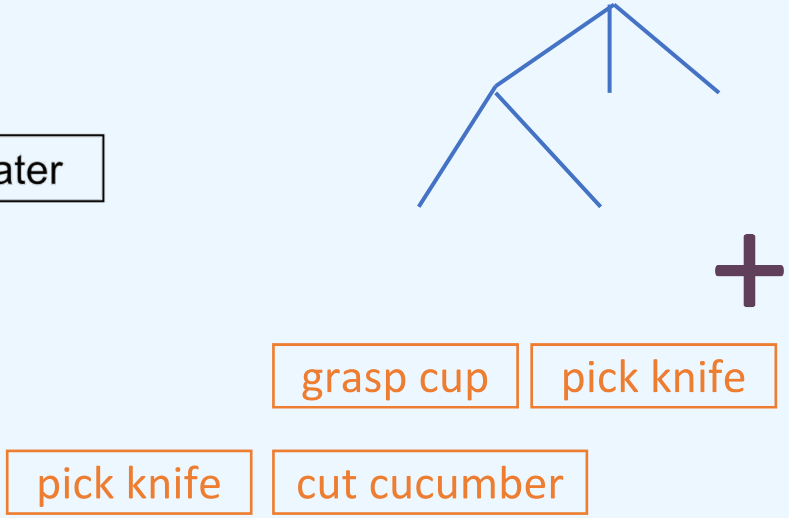
LLM for Planning

- Principle: In the vector space, similar tasks should be close, while distinct tasks should be far away.

Define similarity?

- Example:





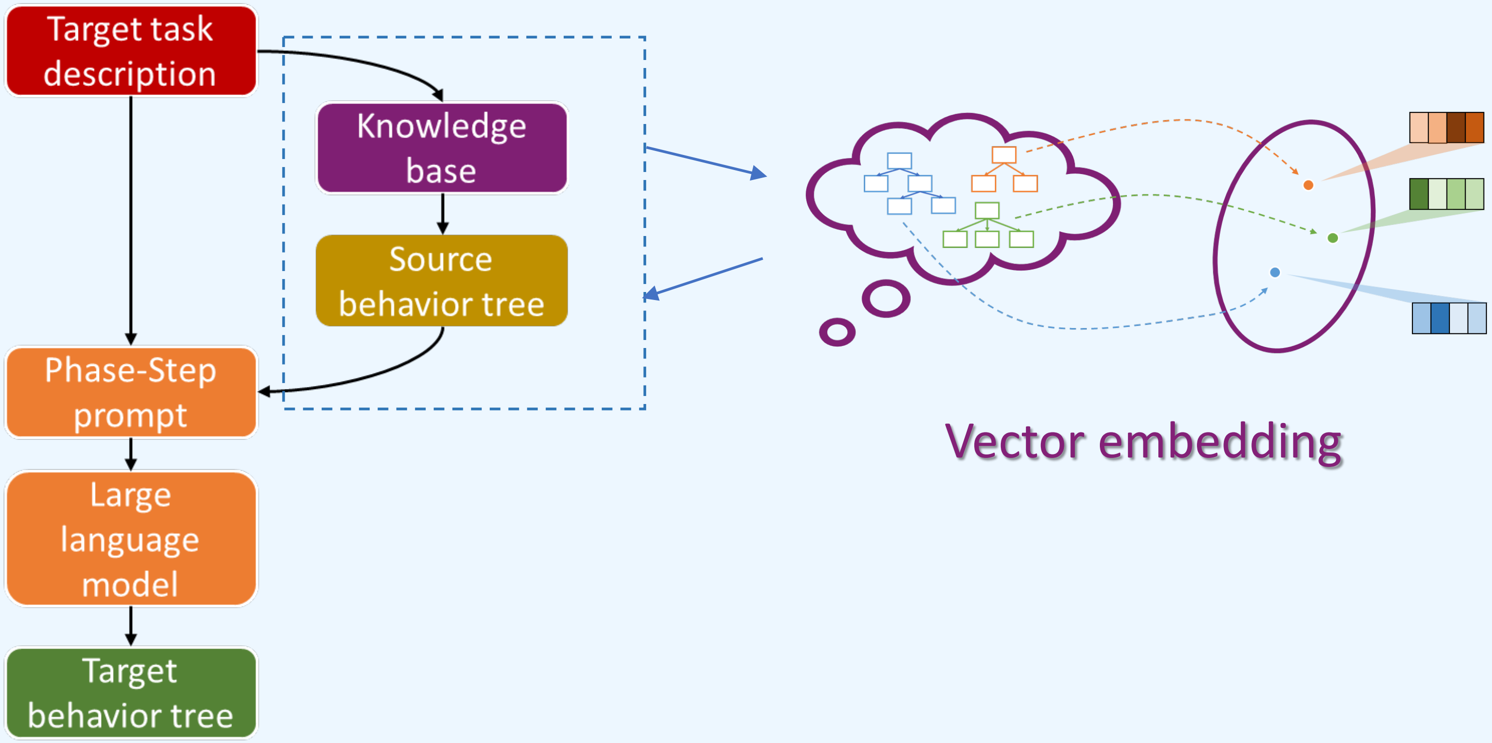
Structural characteristics

+

Semantic characteristics

LLM for Planning

- Summary



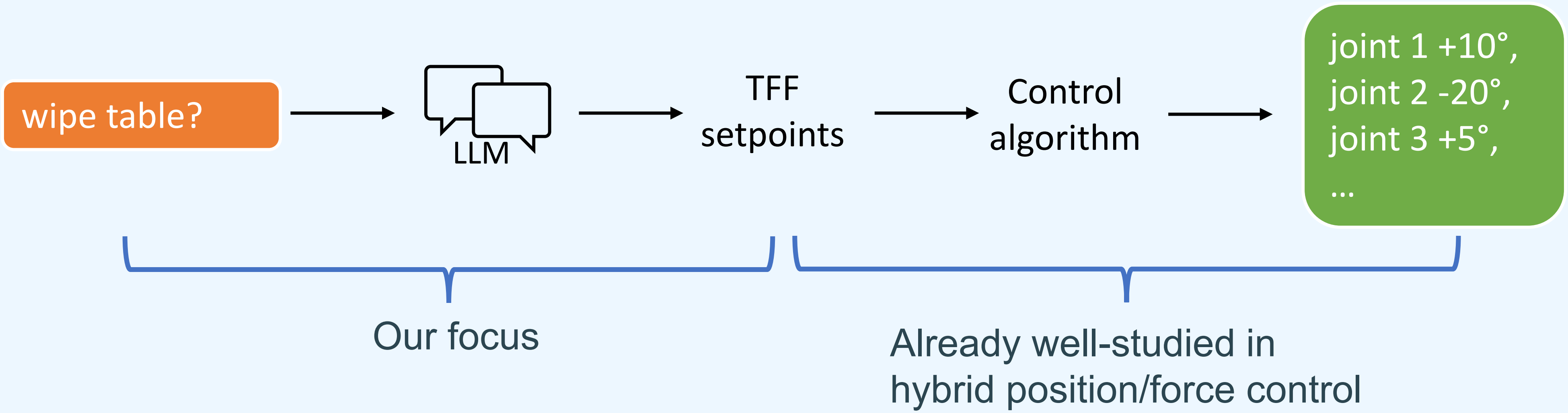
LLM for Acting

- These generated tasks are not yet linked with low-level motor execution.

- Using LLMs to realize this process?

LLM for Acting

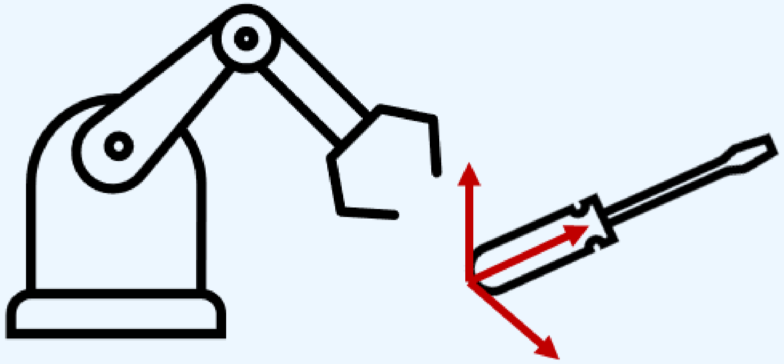
- Scope: Manipulator primitive tasks, typically contact-rich.
- **Input:** a natural-language-described manipulator primitive task.
- **Utilize:** task frame formalism (TFF)
- **Output:** a set of position/force set-points in the task frame.



Yue Cao and C. S. George Lee, “Ground Manipulator Primitive Tasks to Executable Actions using Large Language Models”, AAAI Fall Symposium, 2023

LLM for Acting

- task frame formalism
- A frame specified on the manipulated object,
  3 translational directions,
  3 rotational directions,
  either position controlled/force controlled.



wipe table? → LLM → TFF setpoints → Control algorithm → joint 1 +10°, joint 2 -20°, joint 3 +5°, …

Our focus

Already well-studied in hybrid position/force control

LLM for Acting

The effector velocity and effector force are represented as column vectors in a six-dimensional vector space over the reals:

$$\boldsymbol{v} = (v_x v_y v_z w_x w_y w_z)^T$$

$$\boldsymbol{f} = (f_x f_y f_z g_x g_y g_z)^T$$

Captured from Mason's paper

Structured form! → LLMs prefer structured input, think about the example of generating fake citations.

wipe table? → LLM → TFF setpoints → Control algorithm → joint 1 +10°, joint 2 -20°, joint 3 +5°, …

Our focus

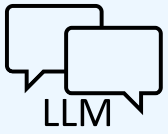
Already well-studied in hybrid position/force control

Mason, M. T., " Compliance and Force Control for Computer Controlled Manipulators ", IEEE Trans. Syst. Man Cybern, 1981.

LLM for Acting

```python
# Source function 1
def turn_screw(translational_x, translational_y, translational_z,
               angular_x, angular_y, angular_z):
  # Coordinate setting: Z axis as the direction of screw
  translational_x=0
  translational_y=0
  translational_z=-5 # N

  angular_x=0
  angular_y=0
  angular_z=5 # rad/sec
  return translational_x, translational_y, translational_z, \
  angular_x, angular_y, angular_z

# Source function 2
def wipe_table(...)
  ...
  return

# Source function 3
def open_door_from_doorknob(...)
  ...
  return

# Target function
def turn_steering_wheel(translational_x, translational_y, translational_z,
                        angular_x, angular_y, angular_z):
```

A 3-shot prompt example




```python
# Coordinate setting: Z axis as the direction of the steering wheel
translational_x=0
translational_y=0
translational_z=0

angular_x=0
angular_y=0
angular_z=5 # rad/sec
return translational_x, translational_y, translational_z, \
angular_x, angular_y, angular_z
```

LLM Output

- Program-function-like prompt
- Task-frame-formalism-based representation
- Few-shot inference

Overview

Part 1. Research Talk

1. Real Autonomous Robots
2. LLM for Planning
3. LLM for Acting

Part 2. Vision Talk

1. AI for ECE
2. Teaching and Research

AI for ECE

- AI is changing the job market of software engineering.

CNBC MARKETS BUSINESS INVESTING TECH POLITICS CNBC TV INVESTING CLUB PRO

AI hiring frenzy to fuel layoffs in other tech segments as firms strive to balance costs

PUBLISHED THU, JAN 25 2024·6:07 PM EST | UPDATED THU, JAN 25 2024·8:23 PM EST

yahoo!finance Search for news, symbols or companies

Finance Watchlists My Portfolio Markets News Videos Yahoo Finance Plus Screeners Persona

CBS NEWS

Google CEO warns of more layoffs in 2024 amid AI push

AXIOS

Big tech companies are pointing to a renewed focus on AI as a justification for layoffs.

Why it matters: The new trend in tech company layoffs could be less about replacing workers with AI, and more about replacing workers with a smaller number of workers who are more skilled in AI, for now.

Lesson learned

AI for ECE

- Will AI change the job market of ECE?
  How do we prepare for this?

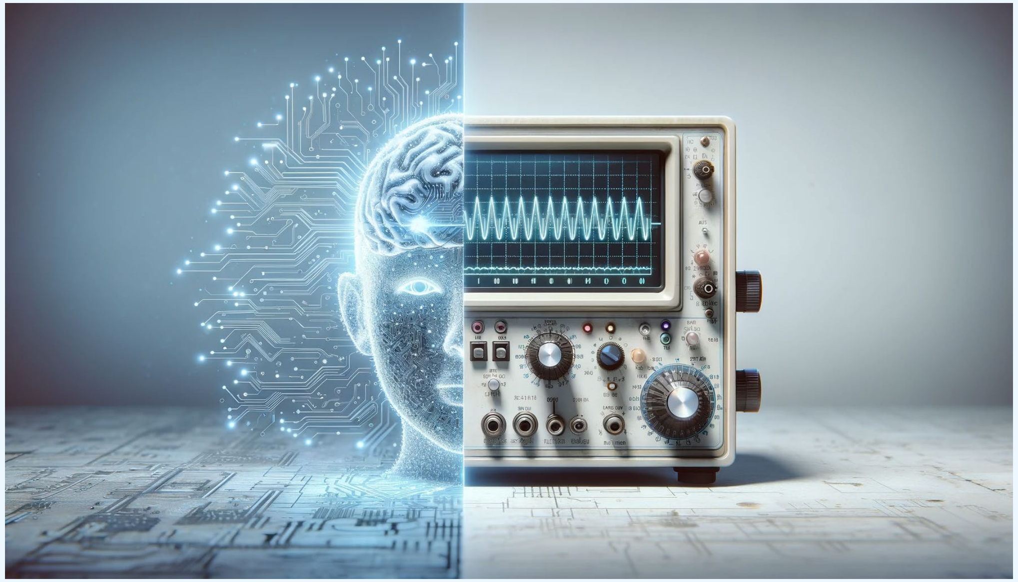
My answer:
Make our ECE students skilled with AI

AI for ECE

- How skilled with AI as an ECE person?

My answer:

As skilled as using Oscilloscope



- We need to know how to use scope as a tool, but not necessarily need to know how to build it from electronic components.

- Similarly for AI.
  We need to know how to apply AI as a tool,
  but not necessarily need to
  1) derive every theorem as MATH
  2) program every code as CS

Picture generated by GPT-4 with DALL·E

Future courses – Robotics/AI

1. Introduction to Robotics

[Soph.]

Basics of various robot systems, manipulator, wheeled/mobile robot/aerial robot, etc…

2. AI Programming with Python

[Soph.]

Basic AI algorithms, with hand-on implementation in Python package, scikit-learn, PyTorch

3. AI in Robotics

[Jr./Sr.]

From symbolic AI reasoning to classical machine learning, also recent advanced Generative AI

4. Robot Simulation

[Jr./Sr. Lab]

Usage of physics-engine-based simulator, such as CoppeliaSim, or GPU-accelerated Nvidia Issac

Future courses – Robotics/AI

5. Applied Deep Learning
[Jr./Sr. /Grad.]
Deep neural networks, generative adversarial network, large language models, …

6. Robot Planning
[Jr./Sr./Grad.]
Pre: ECE 233 Digital Systems
Task planning,
Motion planning

7. Robot Manipulation and Locomotion [Sr./Grad.]
Pre: ECE 320 Control Systems
Manipulator kinematics and dynamics,
Legged robot locomotion

8. Autonomous Driving
[Grad.]
Pre: ECE 425 Mobile Robots
Vehicle dynamics, map system, sensor fusion, driver intent, …
hand-on project with dSPACE RTMaps

Future courses – Robotics/AI

And most importantly,
the new courses that students want.

Undergraduate research– Robotics/AI

- Understand students’ work load and specialty
- Research idea matters!
- Present our work outside.
  Not as regular paper,
  but as poster/short report/workshop paper…

Undergraduate research– Robotics/AI

Regional events

MMLS 2023
OVERVIEW SPEAKERS SCHEDULE POSTERS SPONSORS CODE OF CONDUCT REGISTER

Machine Learning Research in the Midwest

Washington University in St. Louis

8th Midwest Workshop on Control and Game Theory

HOME HOTEL & TRAVEL POSTER ABSTRACTS SPEAKERS REGISTRATION PROGRAM

International events

IEEE Robotics & Automation Society

AAAI
Association for the Advancement of Artificial Intelligence

Overview

Part 1. Research Talk

1. Real Autonomous Robots
2. LLM for Planning
3. LLM for Acting

Part 2. Vision Talk

1. AI for ECE
2. Teaching and Research

Thanks!

Yue Cao, yuecao@purdue.edu
Purdue ECE