

Final Project

COMP7607A



TA Introduction (COMP7607A)

Gao Peifeng

2024-present:

PhD, The University of Hong Kong

Research Interest:

Deep Learning Theory

E-mail: gaopeifeng@connect.hku.hk

Wang Jiannan

2023-present:

PhD, The University of Hong Kong

Research Interest:

Machine Learning System

E-mail: steaunk@connect.hku.hk

Final Project Announcement

Teams: 5-6 students

Topics:

Creating and Developing LLM Agents for Applications.

Any Real-world applications that can utilize AI agent is ok.

What is LLM (Large Language Model)

$$answer = LLM(question, context)$$

What it can do

$$answer = LLM(question, context)$$

What is LLM

Can answer questions.

*question: what's the main idea
Context: paper*



LLM



Answer: the summaried idea of the given paper

What is LLM

$$\textit{answer} = \textit{LLM}(\textit{question}, \textit{context})$$

Can answer questions.

But far from JARVIS in the iron man.

How make it Order a takeout meal?

What is LLM

answer = LLM(question, context)



function calling = LLM(order, context)

function calling = LLM(order, context)

What is LLM

Python function: *answer = order_a_takeout(meal, address)*

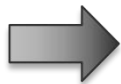
Order(question):

- Order me some takeout, I live in..., and I'd like to eat ...
- Return your answer as function call (*order_a_takeout*)

Context: The restaurants around my home



LLM

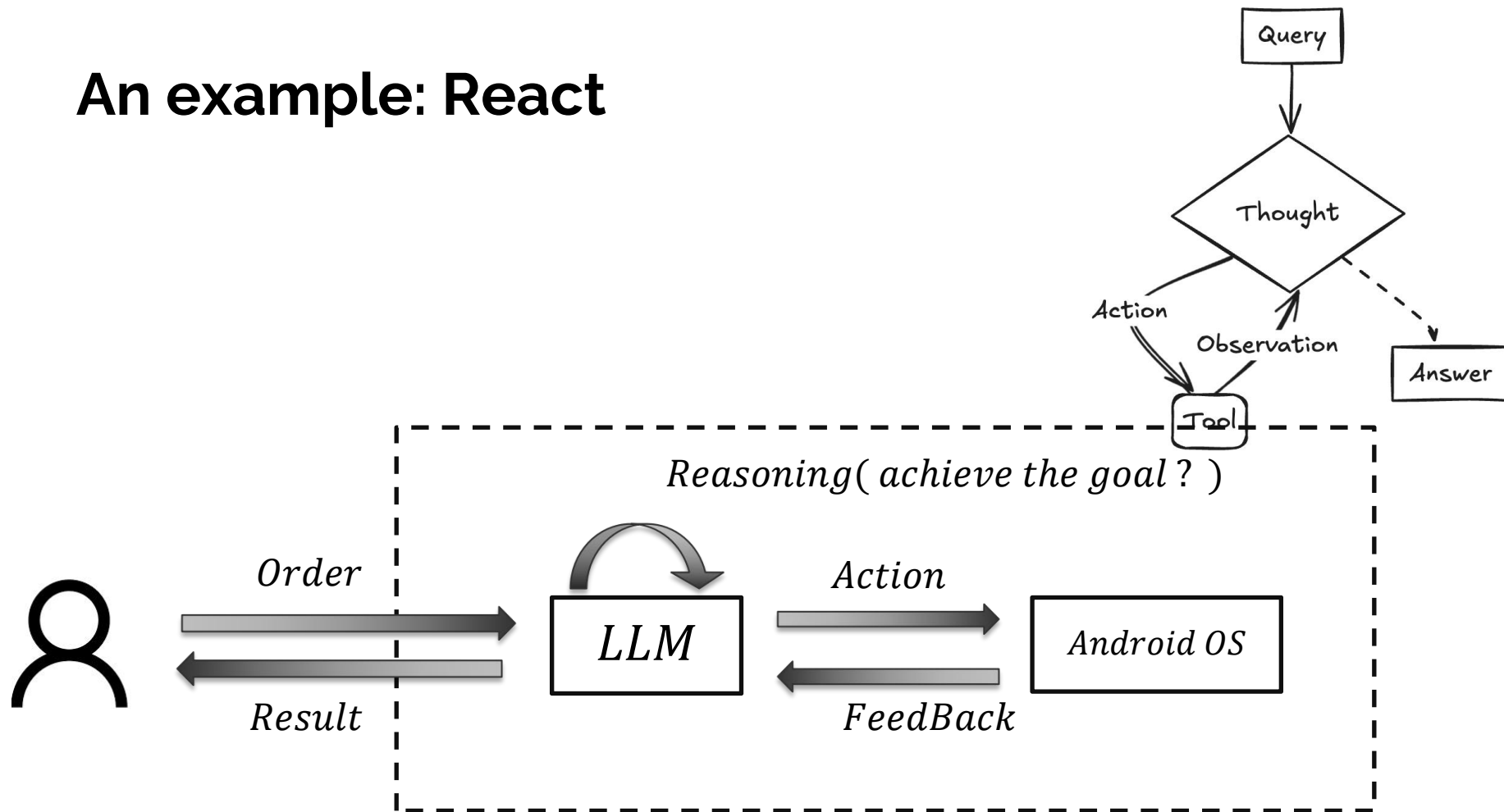


function_call(Answer):
order_a_takeout(meal, address)

A demo: Find good to eat around on Android

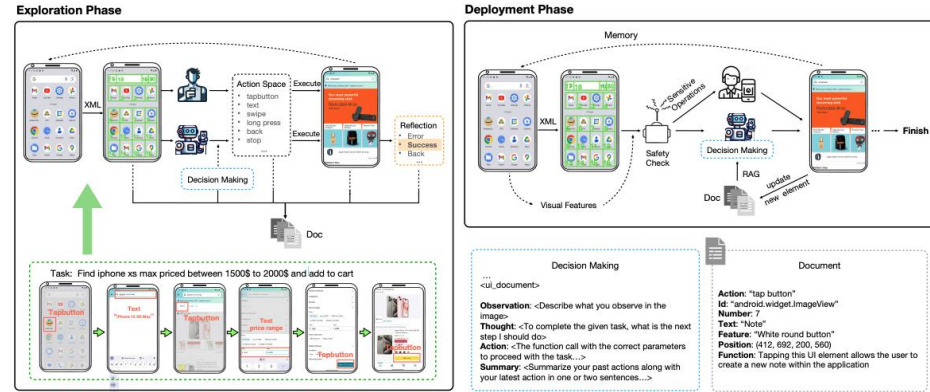


An example: React



Exemplary topic 1: Mobile Agents

- **AppAgent**
 - Exploration & deployment
- **AppAgent v2**
 - Structured storage format with RAG
- **AppAgentX**
 - Evolutionary memory mechanism

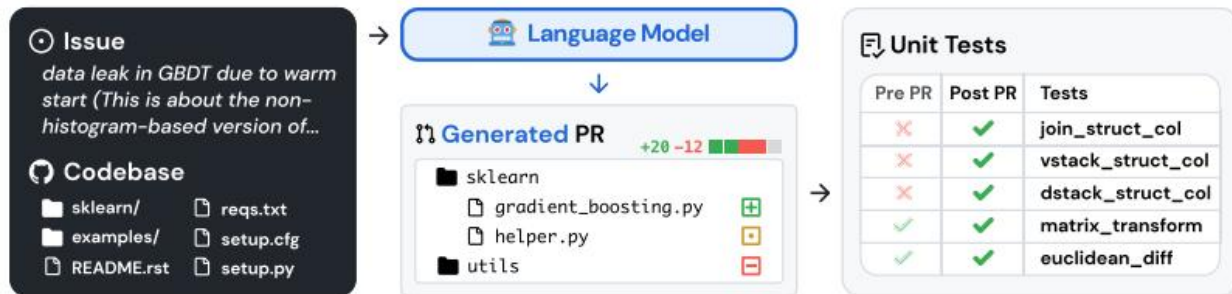


[1] AppCopilot: Toward General, Accurate, Long- Horizon, and Efficient Mobile Agent.

[2] AppAgent: Multimodal Agents as Smartphone Users.

[3] AppAgent v2: Advanced Agent for Flexible Mobile Interactions

Exemplary topic 2: Coding Agent

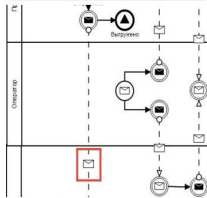


Diagramming

Show message element name

Currently, names of message elements on message flows are not rendered

Given this example diagram [Image] ...



bpmn-js

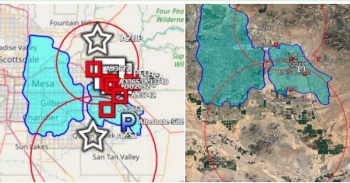
Interactive Mapping

KML Symbol Align/Placement/Size

There is a bug with the anchor point for some symbols

I've attached a screen clipping from Google Earth to show how it is supposed to look.

[Right Image] ...



openlayers

Syntax Highlighting

Bracket highlighted with different color in class inheritance context.

~ Reproduced in JSFiddle: <https://jsfiddle.net/kkangmj/e7h48w36/7/> [Image] ...

```
open class Tag
class TABLE: Tag {
  fun tr(init: TR.() -> Unit)
}
class TR: Tag {
  fun td(init: TD.() -> Unit)
}
class TD: Tag
```

highlight.js

Web Frameworks

[CascaderSelect]使用虚拟滚动时背景色异常

Component: CascaderSelect
Steps to reproduce
[Image] ...



next

[1] SWE-bench Multimodal: Do AI Systems Generalize to Visual Software Domains?

[2] SWE-bench: Can Language Models Resolve Real-World GitHub Issues?

[3] SWE-RL: Advancing LLM Reasoning via Reinforcement Learning on Open Software Evolution

[4] SwingArena: Competitive Programming Arena for Long-context GitHub Issue Solving

Details

1. **Project proposal (10%):**

Up to 2 pages (exclude references)

proposal containing team information and a brief introduction of the project

2. **Implementation, Presentation and demonstration (60%):**

Present your project in the class and showcase your exciting demos!

Evaluated based on:

- a. Technical contribution (30%)
- b. Excitement (15%), soundness (10%)
- c. Presentation clearness (5%)

3. **Project Report (30%):**

4-8 pages report concludes the project.

Reference

- [1] AppCopilot: Toward General, Accurate, Long- Horizon, and Efficient Mobile Agent.
- [2] AppAgent: Multimodal Agents as Smartphone Users.
- [3] AppAgent v2: Advanced Agent for Flexible Mobile Interactions
- [4] AppAgentX: Evolving GUI Agents as Proficient Smartphone Users
- [5] MobileSteward: Integrating Multiple App-Oriented Agents with Self-Evolution to Automate Cross-App Instructions
- [6] Mobile-Agent-v2: Mobile Device Operation Assistant with Effective Navigation via Multi-Agent Collaboration
- [7] A3: Android Agent Arena for Mobile GUI Agents
- [8] MMBench-GUI: Hierarchical Multi-Platform Evaluation Framework for GUI Agents
- [9] SWE-bench Multimodal: Do AI Systems Generalize to Visual Software Domains?
- [10] SWE-bench: Can Language Models Resolve Real-World GitHub Issues?
- [11] SWE-RL: Advancing LLM Reasoning via Reinforcement Learning on Open Software Evolution
- [12] SwingArena: Competitive Programming Arena for Long-context GitHub Issue Solving
- [13] ADK: An Agent Development Kit Based on a Formal Design Model for Multi-Agent Systems
- [14] Agentic AI for Intent-Based Industrial Automation
- [15] Autonomous Control Leveraging LLMs: An Agentic Framework for Next-Generation Industrial Automation
- [16] The Rise and Potential of Large Language Model Based Agents: A Survey
- [17] Large Language Model Agent: A Survey on Methodology, Applications and Challenges