Large Language Model Powered Agents in the Web

Tutorial at The Web Conference 2024 in Singapore (WWW 2024)

Yang Deng[1], An Zhang[1], Yankai Lin[2], Xu Chen[2], Ji-Rong Wen[2], Tat-Seng Chua[1]

[1] NExT++ Research Centre, National University of Singapore
[2] Gaoling School of Artificial Intelligence, Renmin University of China

dengyang17dydy@gmail.com, an_zhang@nus.edu.sg, yankailin@ruc.edu.cn
xu.chen@ruc.edu.cn, jrwen@ruc.edu.cn, chuats@comp.nus.edu.sg

May 13, 2024, Singapore

Speakers

Yang Deng
NUS National University of Singapore

An Zhang
NUS National University of Singapore

Yankai Lin
Renmin University of China

Xu Chen
Renmin University of China

Jirong Wen
Renmin University of China

Tat-Seng Chua
NUS National University of Singapore

Outline

- **Part 1: Introduction of LLM-powered Agents**

- Part 2: LLM-powered Agents with Tool Learning

- Part 3: LLM-powered Agents in Social Network

- Part 4: LLM-powered Agents in Recommendation

- Part 5: LLM-powered Conversational Agents

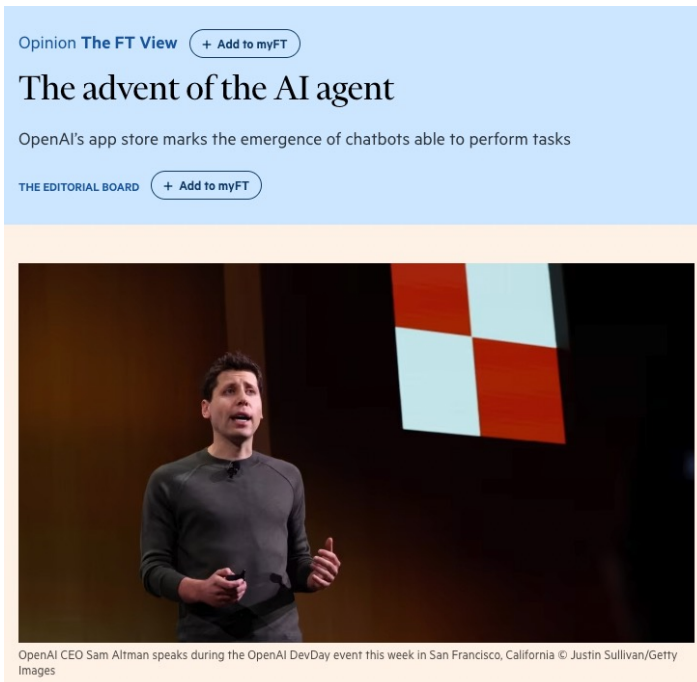- Part 6: Open Challenges and Beyond

Aim of AGI

- **Large LLMs** exhibit characteristics of **artificial general intelligence (AGI)**, which has **cognitive abilities** similar to that of human.

- In other words, AI can now perform most functions that humans are capable of doing.

Linguistic Intelligence → *Towards AGI* → **Artificial General Intelligence**

Observation | Remember | Reasoning | Using tools | Action | Learn from others | Self-reflection | Autonomous nature

AI Agents

- **LLM-powered Agents** are artificial entities that **enhance LLMs** with **essential capabilities**, enabling them to sense their environment, make decisions, and take actions.

Opinion **The FT View** + Add to myFT

The advent of the AI agent

OpenAI's app store marks the emergence of chatbots able to perform tasks

THE EDITORIAL BOARD + Add to myFT

OpenAI CEO Sam Altman speaks during the OpenAI DevDay event this week in San Francisco, California © Justin Sullivan/Getty Images

- **Sam Altman** (Former CEO of OpenAI) himself said in his keynote: *"GPTs and Assistants are **precursors** to **agents**. They will gradually be able to plan and to perform more complex actions on your behalf. These are our first step toward AI Agents."*

- **Bill Gates** said in his BLOG: *"**Agents** are not only going to change how everyone interacts with computers. They're also going to **upend the software industry**, bringing about the biggest revolution in computing since we went from typing commands to tapping on icons."*

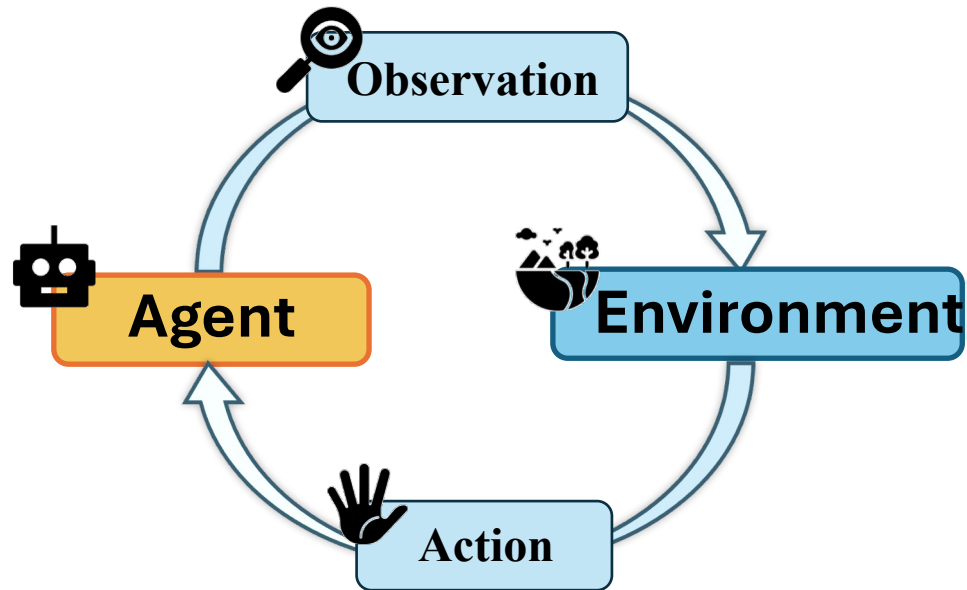 **be my eyes** **AI-powered visual assistance.**

- Application:

News in Financial Times. " The advent of the AI agent" .
GatesNotes. "The Future of Agents: AI is about to completely change how you use computers".

■ This paves the way for the use of AI agents to simulate users and other entities, as well as their interactions.

Observation

Agent

Environment

Action

Environment

➤ The external **context** or **surroundings** in which the agent operates and makes decisions.

• Human & Agents' behaviors

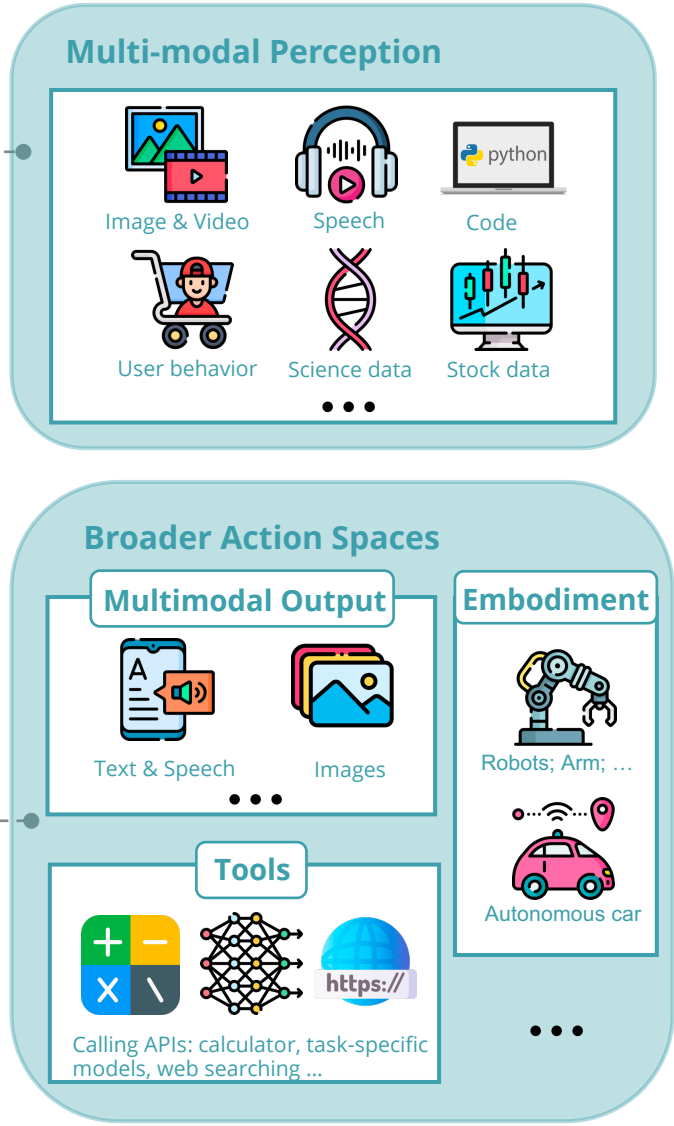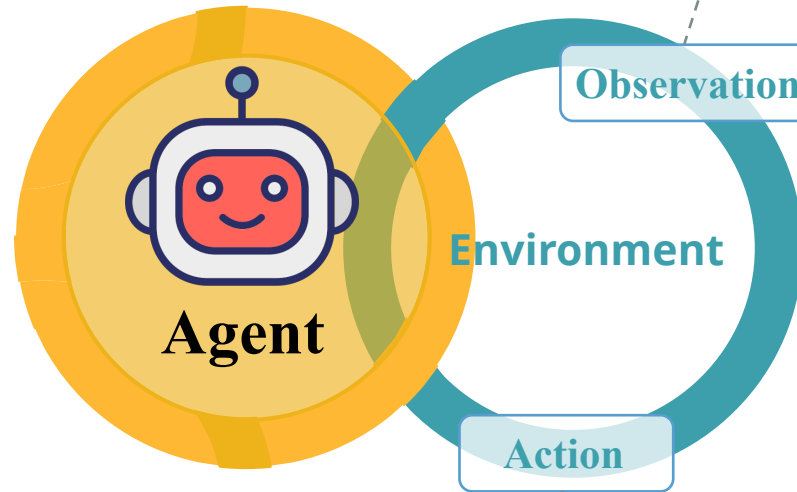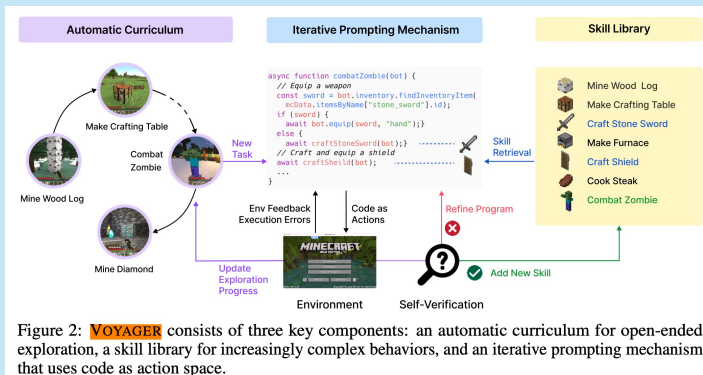• External database and knowledges
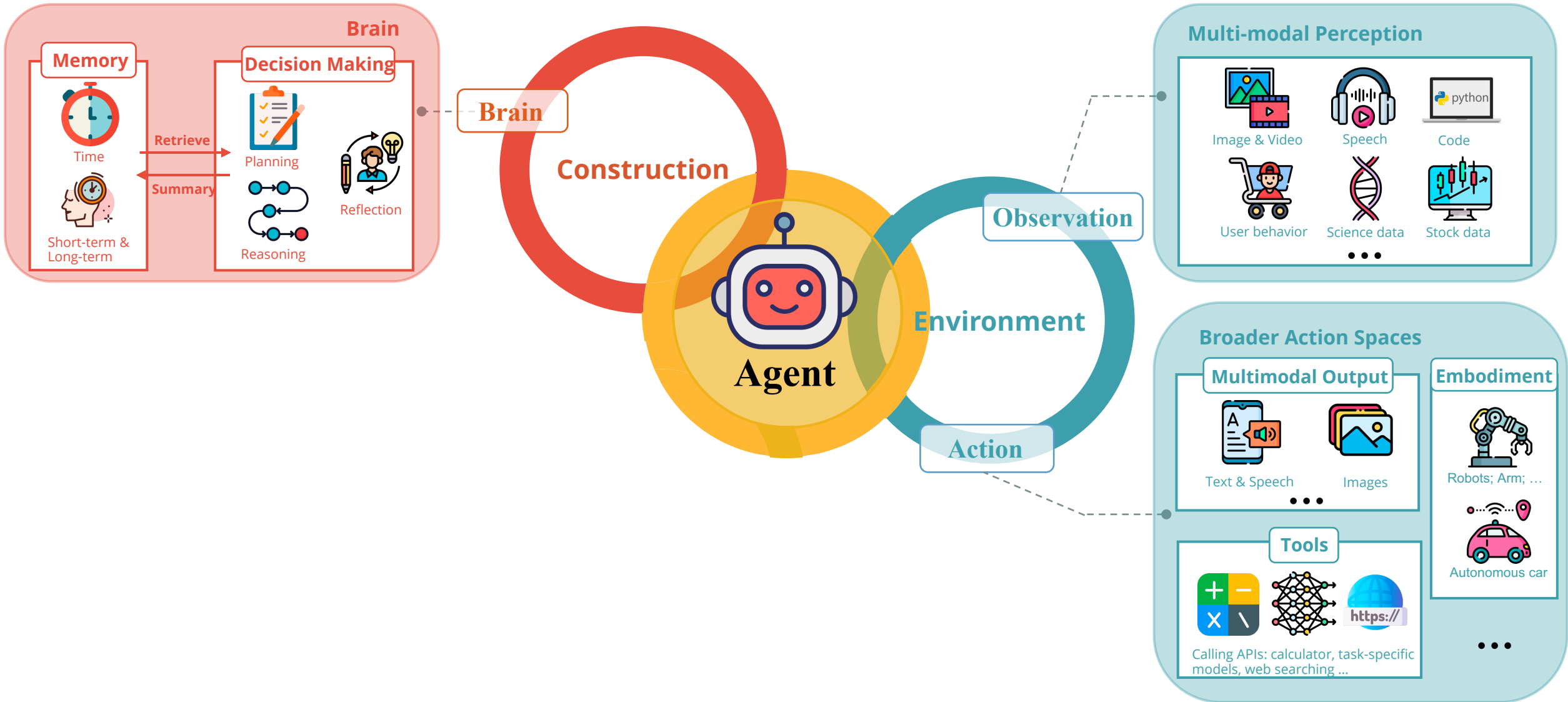
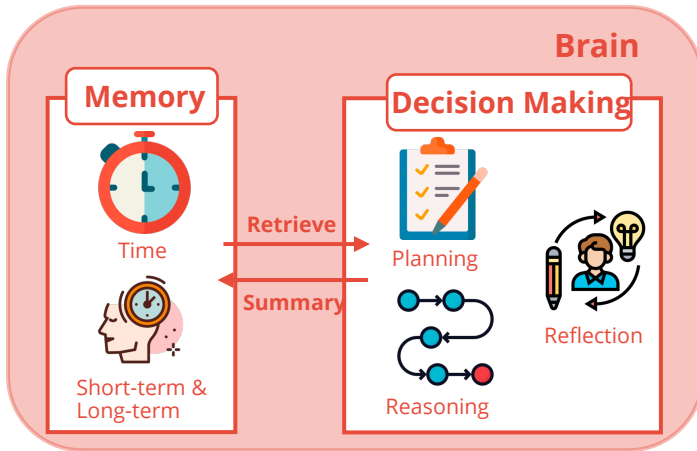

• Virtual & Physical environment

NEXT++

Action

➤ call external **APIs** for extra information that is missing from the model weights (often hard to change after pre-training): **Generating multimodal outputs; Embodied Action; Learning tools; Using tools; Making tools; ......**

Figure 2: VOYAGER consists of three key components: an automatic curriculum for open-ended exploration, a skill library for increasingly complex behaviors, and an iterative prompting mechanism that uses code as action space.

Agent

Observation

Environment

Action

Multi-modal Perception

Image & Video   Speech   Code

User behavior   Science data   Stock data

· · ·

Broader Action Spaces

Multimodal Output

Text & Speech   Images

· · ·

Embodiment

Robots; Arm; ...

Autonomous car

Tools

Calling APIs: calculator, task-specific models, web searching ...

· · ·

Guanzhi Wang et al., Voyager: An Open-Ended Embodied Agent with Large Language Models.

- ❑ **Memory:** "memory stream" stores sequences of agent's past observations, thoughts and actions:
  - ➢ Sufficient space for long-term and short-term memory;
  - ➢ Abstraction of long-term memory;
  - ➢ Retrieval of past relevant memory;

- ❑ **Decision Making Process:**
  - ➢ **Planning: Subgoal** and **decomposition**: Able to break down large tasks into smaller, manageable subgoals, enabling efficient handling of complex tasks.
  - ➢ **Reasoning:** Capable of doing **self-criticism** and **self-reflection** over past actions**, learn from mistakes** and **refine** them for future steps, thereby improving the quality of final results.
- ❑ Personalized memory and reasoning process foster **diversity** and **independence** of AI Agents.