

AI 大模型开发工程师 LangChain全面剖析之Agents

讲师：李希沅

目录

- 1 Agent的设计理念
- 2 Chains设计的优点和缺点
- 3 LangChain的Agents模块设计理念剖析
- 4 LangChain提供的Agents抽象

1 Agent的设计理念

01、Agent的设计理念



Andrew Ng



@AndrewYNg

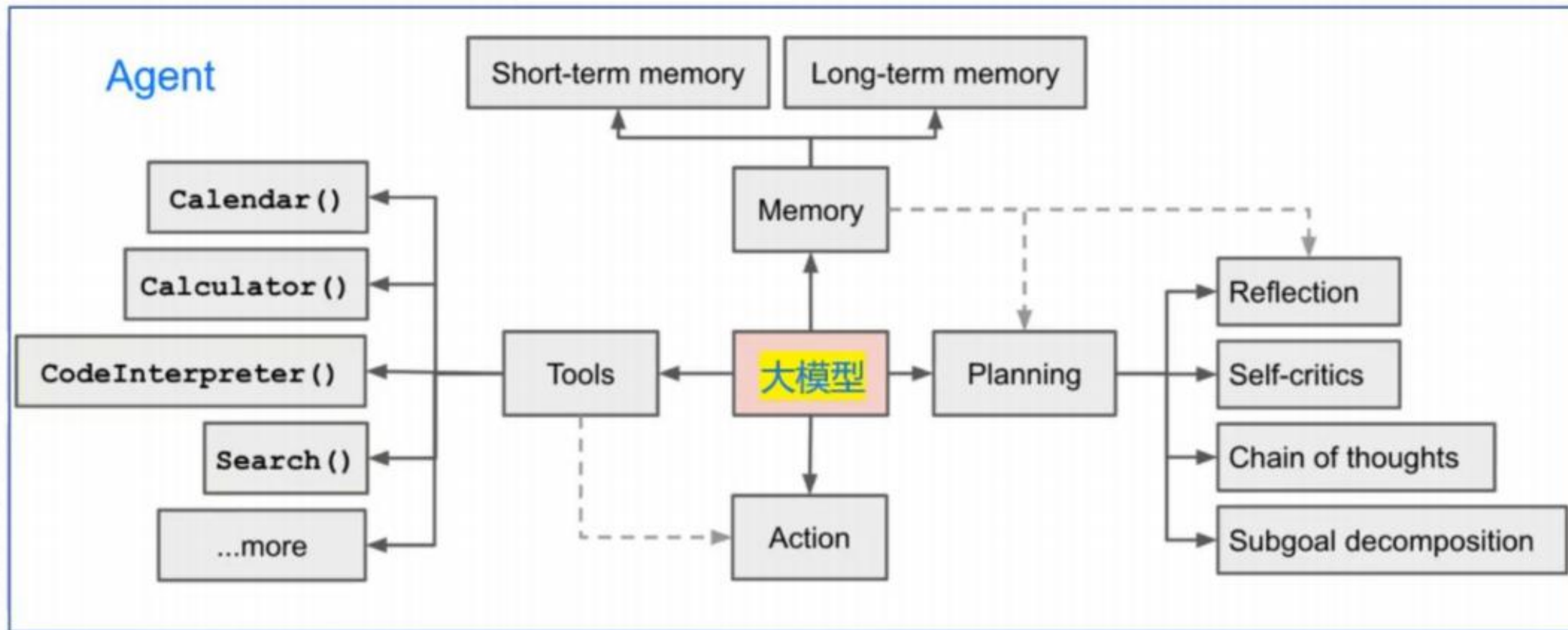


I think AI agentic workflows will drive massive AI progress this year — perhaps even more than the next generation of foundation models. This is an important trend, and I urge everyone who works in AI to pay attention to it.

吴恩达教授提出：当前，我们主要在零样本模式下使用 LLM，提供 prompt，逐个 token 地生成最终输出，没有进行调整。这类似于要求某人从头到尾写一篇文章，直接打字，不允许退格，并期望得到高质量的结果。尽管有困难，LLM 在这项任务上仍然表现得非常好！然而，通过智能体工作流，我们可以要求 LLM 多次迭代文档。例如，它可能需要执行一系列步骤：

- 规划大纲；
- 决定需要进行哪些网络搜索（如果需要），来收集更多信息；
- 写初稿；
- 通读初稿，找出不合理的论点或无关信息；
- 修改草稿；

02、Agent的设计理念

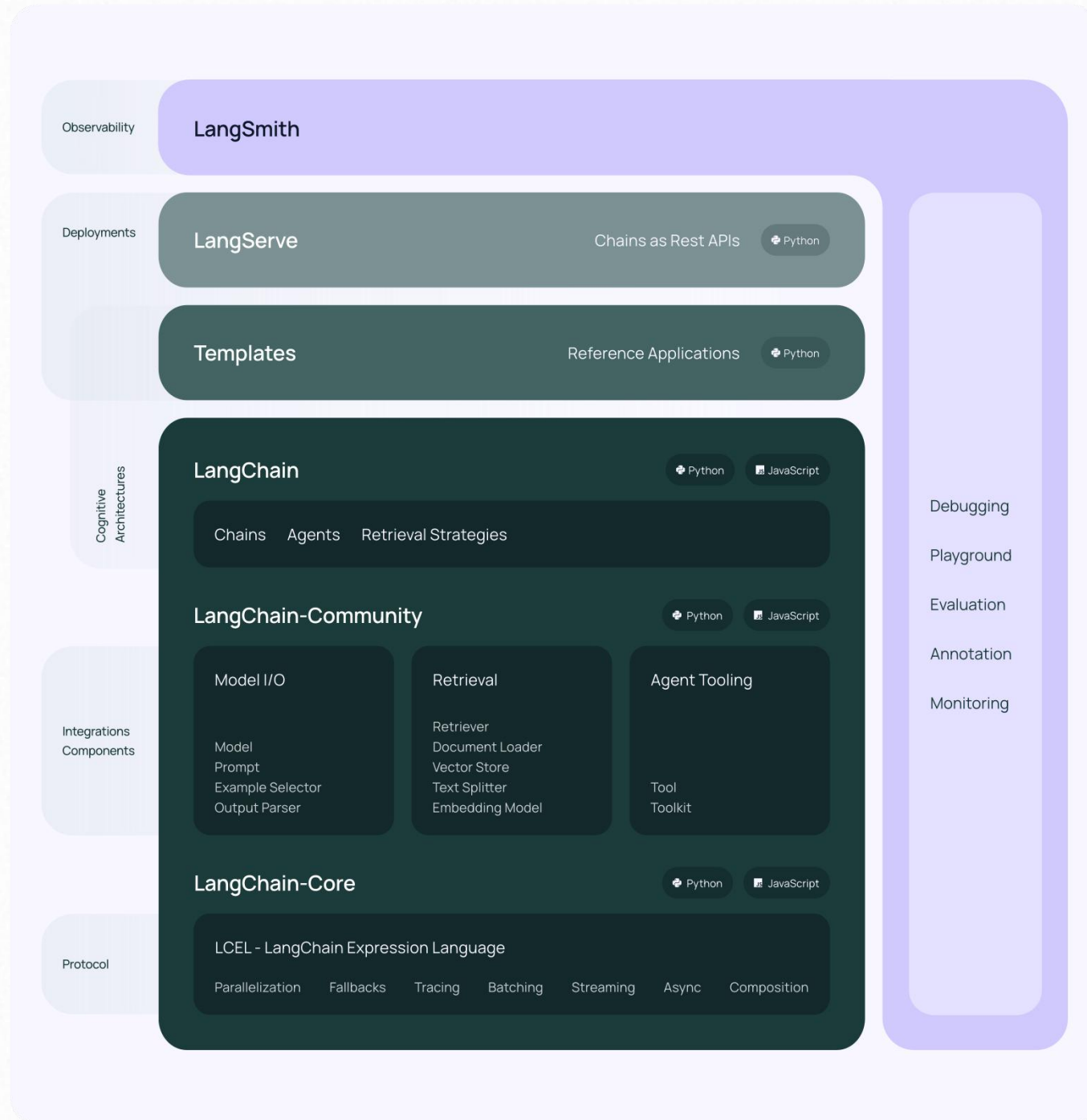


03、LangChain的设计理念

面向LLM开发

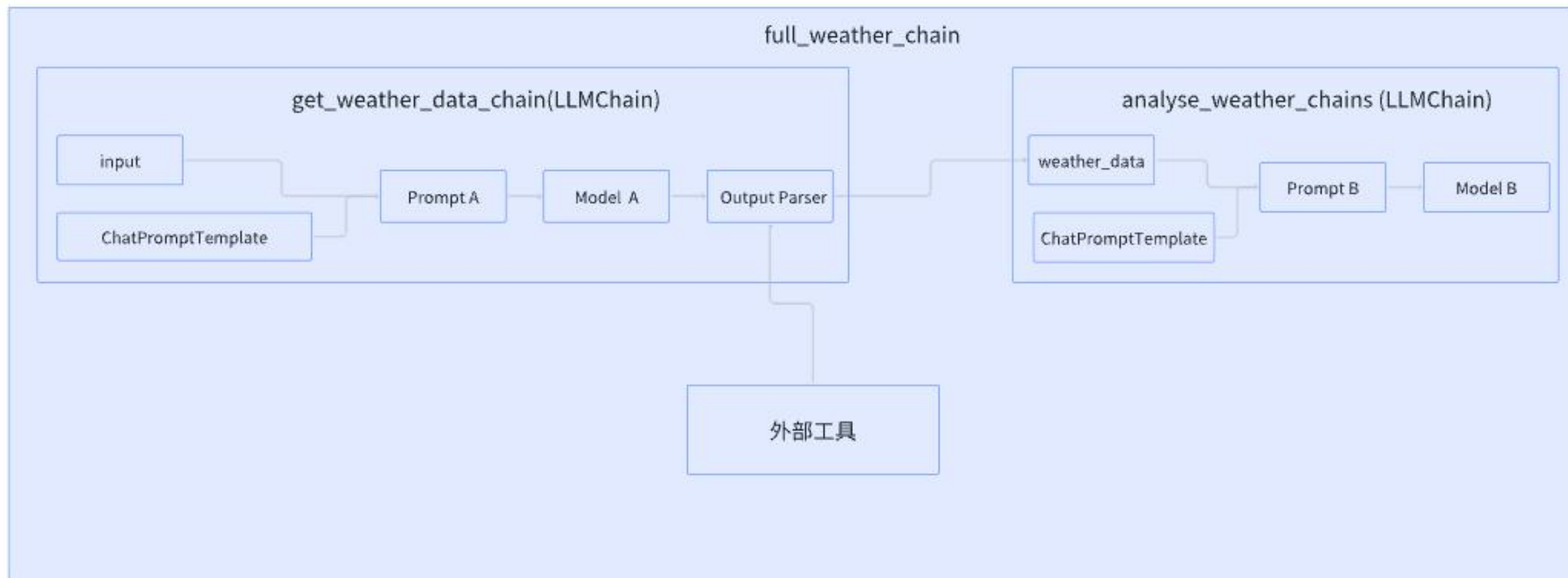
面向Chain开发

面向Agent开发



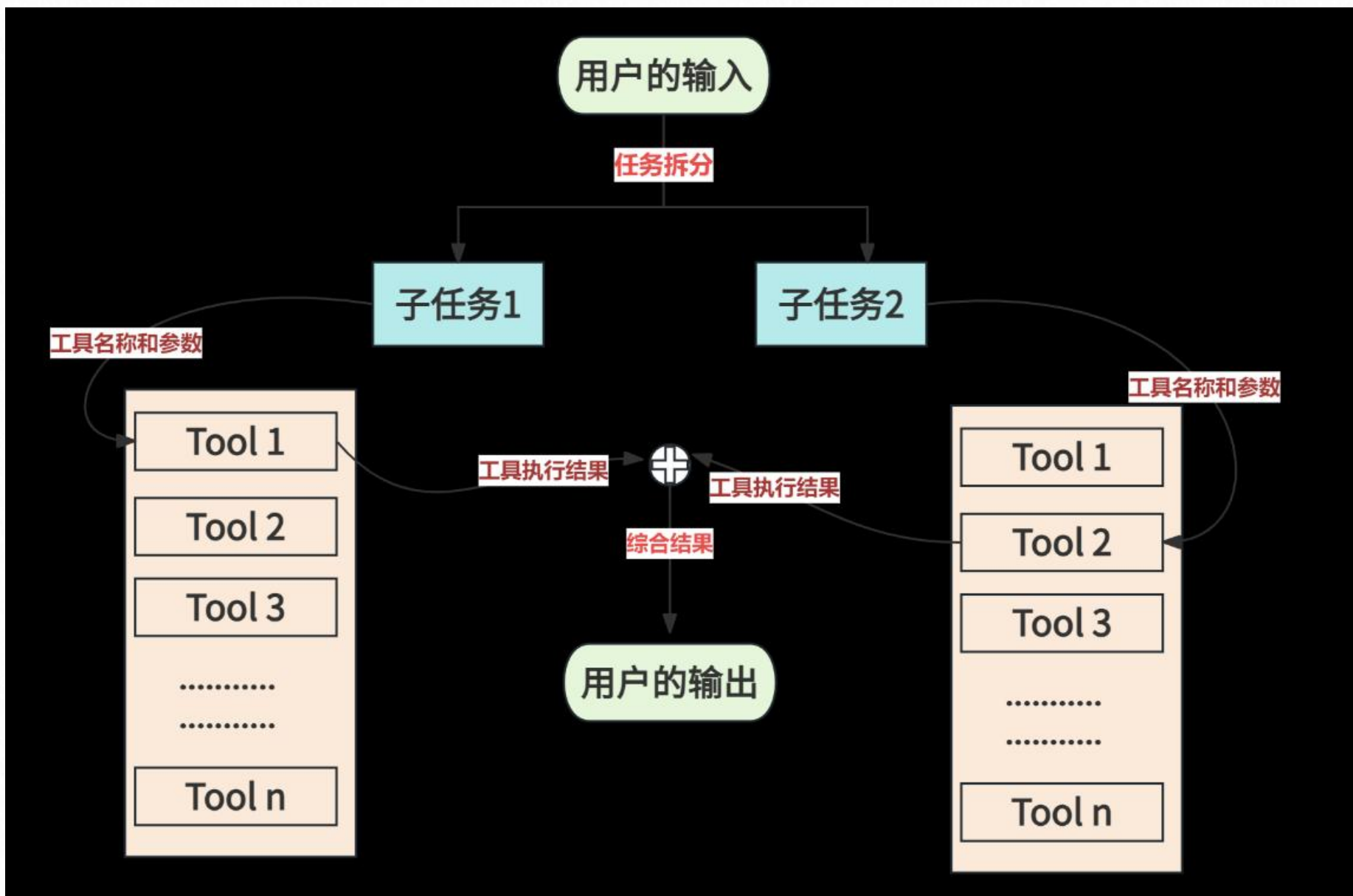
② Chains设计的优点和缺点

01、基于Chains构建获取天气功能



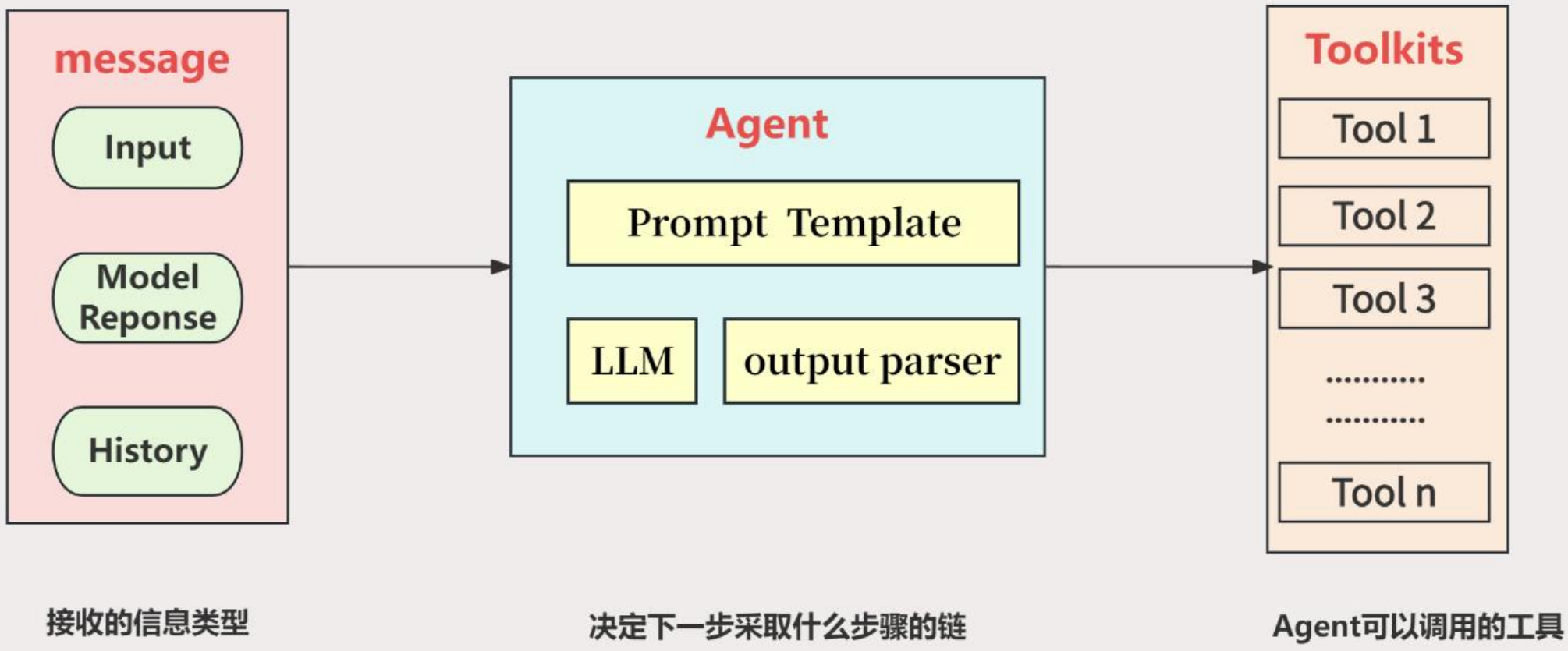
3 LangChain的Agents模块设计理念

01、Agents的设计理念



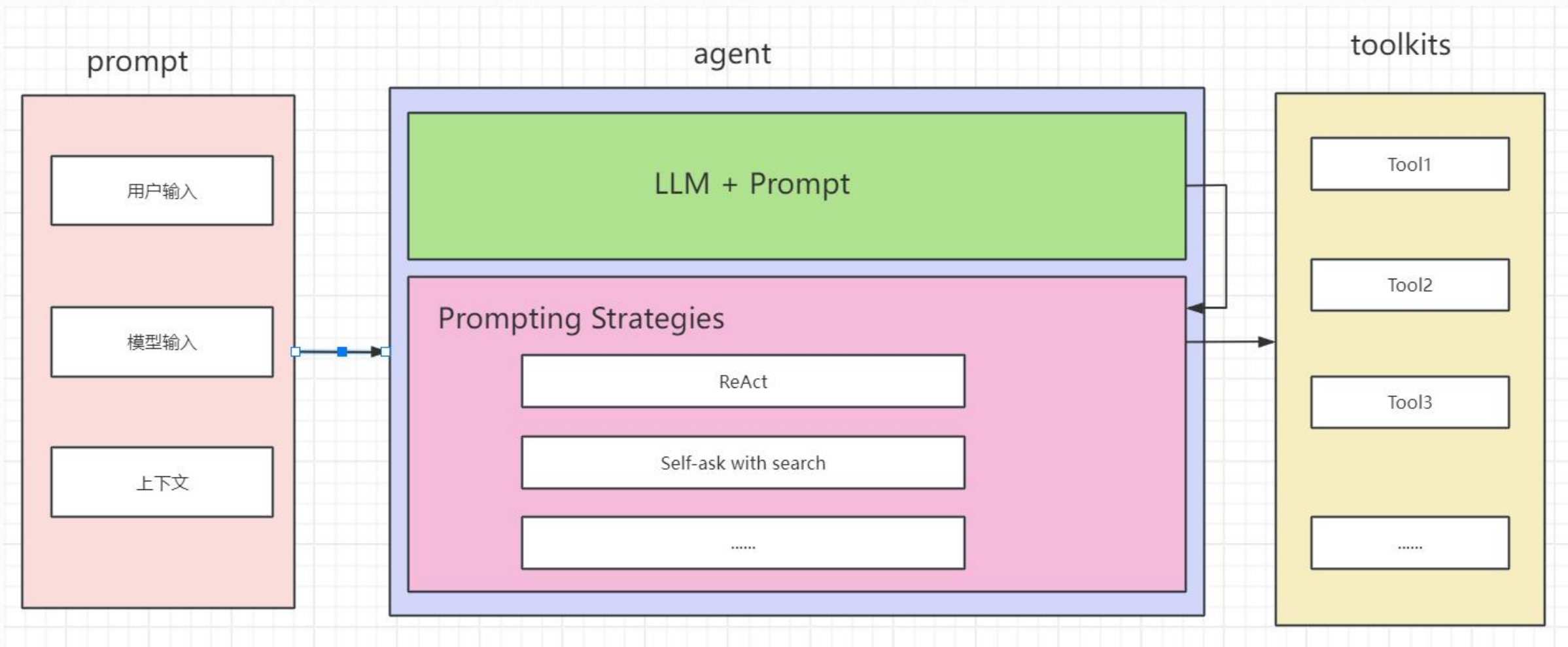
02、Agents的设计理念

AgentExecutor：实际调用Agent，执行它选择的操作，将操作输出传递回Agent，然后重复



01、Agents的设计理念

AgentExecutor

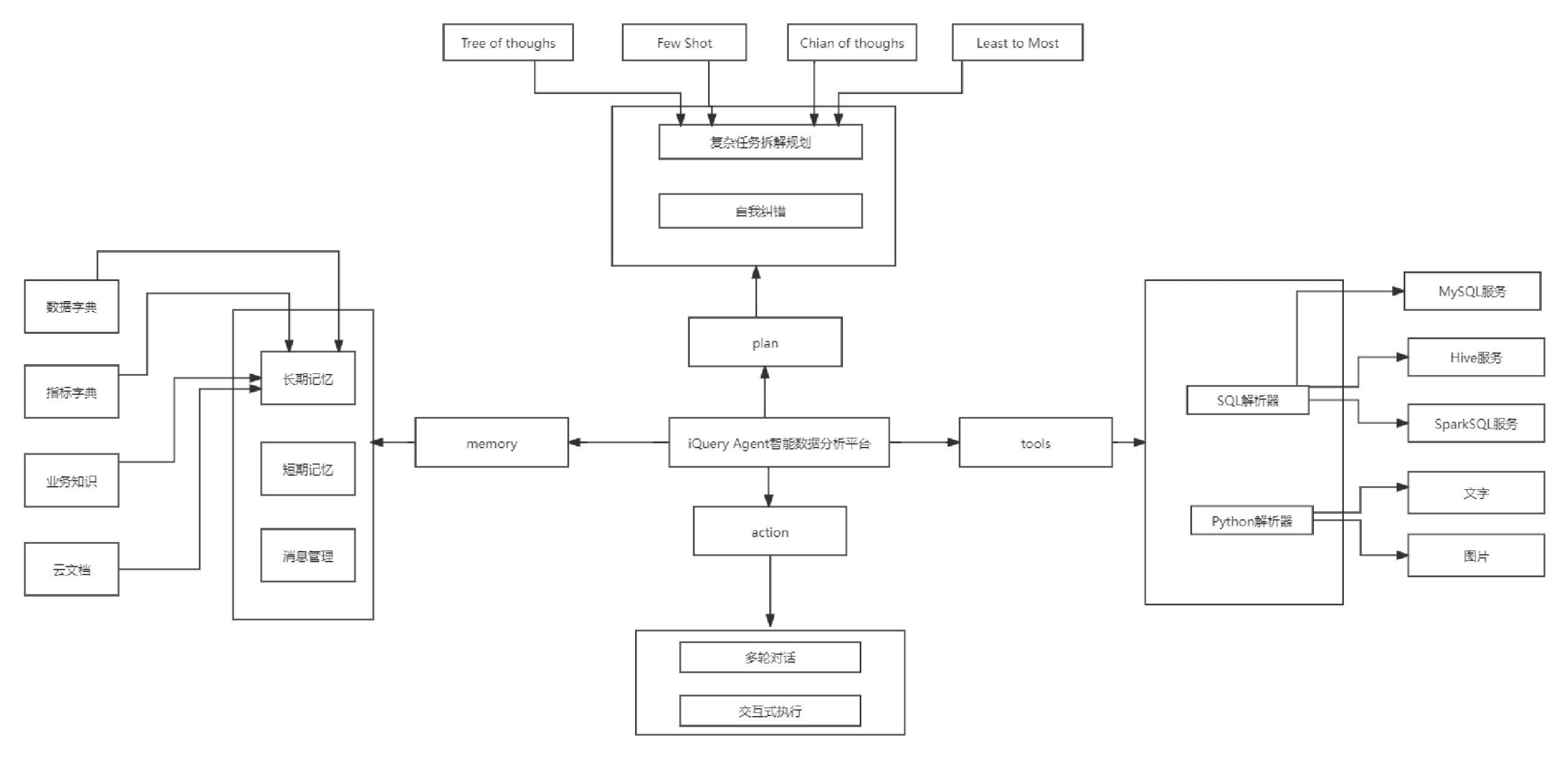


4 LangChain提供的Agents抽象

01、 Agent Type

Agent Type	Intended Model Type	Supports Chat History	Supports Multi-Input Tools	Supports Parallel Function Calling	Required Model Params	When to Use	API
Tool Calling	Chat	✔	✔	✔	tools	If you are using a tool-calling model	Ref
OpenAI Tools	Chat	✔	✔	✔	tools	[Legacy] If you are using a recent OpenAI model (1106 onwards). Generic Tool Calling agent recommended instead.	Ref
OpenAI Functions	Chat	✔	✔		functions	[Legacy] If you are using an OpenAI model, or an open-source model that has been finetuned for function calling and exposes the same functions parameters as OpenAI. Generic Tool Calling agent recommended instead	Ref
XML	LLM	✔				If you are using Anthropic models, or other models good at XML	Ref
Structured Chat	Chat	✔	✔			If you need to support tools with multiple inputs	Ref
JSON Chat	Chat	✔				If you are using a model good at JSON	Ref
ReAct	LLM	✔				If you are using a simple model	Ref
Self Ask With Search	LLM					If you are using a simple model and only have one search tool	Ref

02、智能数据分析平台



03、Tookits

工具名称	中文说明
Apify	用于网络抓取和自动化的工具
ArXiv API Tool	提供访问学术论文和文献的工具
AWS Lambda API	提供云函数或无服务器计算的工具
Shell Tool	允许机器学习模型直接与系统 shell 交互的工具
Bing Search	提供 Bing 搜索引擎查询结果的工具
Brave Search	提供 Brave 搜索引擎查询结果的工具
ChatGPT Plugins	插件工具，用于扩展 ChatGPT 的功能
DataForSeo API Wrapper	从各种搜索引擎获取搜索结果的工具
DuckDuckGo Search	提供 DuckDuckGo 搜索引擎查询结果的工具
File System Tools	提供与本地文件系统交互的工具集
Golden Query	提供基于知识图谱的自然语言查询服务的工具
Google Places	提供 Google 地点信息和查询的工具
Google Search	提供 Google 搜索引擎查询结果的工具
Google Serper API	用于网络搜索的工具，需要注册并获取 API 密钥
Gradio Tools	提供与 Gradio 应用交互的工具
GraphQL tool	提供 GraphQL 查询的工具
huggingface_tools	提供了与 HuggingFace 库交互的工具
Human as a tool	描述了如何将人类作为一个执行具体任务的工具
IFTTT WebHooks	提供与 IFTTT WebHooks 交互的工具

工具名称	中文说明
Lemon AI NLP Workflow Automation	用于自动化 NLP 工作流的工具
Metaphor Search	专为机器学习模型设计的搜索引擎工具
OpenWeatherMap API	提供查询天气信息的工具
PubMed Tool	提供 PubMed 医学文献查询的工具
Requests	提供网络信息获取功能的工具
SceneXplain	提供图像描述服务的工具
Search Tools	提供各种搜索工具的工具集
SearxNG Search API	提供自托管 SearxNG 搜索 API 查询的工具
SerpAPI	提供网络搜索的工具
Twilio	提供 SMS 或其他消息通道发送消息的工具
Wikipedia	提供查询 Wikipedia 百科全书的工具
Wolfram Alpha	提供查询 Wolfram Alpha 的知识引擎的工具
YouTubeSearchTool	提供 YouTube 视频搜索的工具
Zapier Natural Language Actions API	用于自然语言处理的 API 工具

03、ReAct举例说明

帮我去搜索5篇关于RAG的论文

正在行动:

我需要在arxiv上搜索RAG(Reading Comprehension Answer Generator)论文。

Action: arxiv

Action Input: RAG reading comprehension

获得数据:

Published: 2024-02-19

Title: CRUD-RAG: A Comprehensive Chinese B... ...

正在行动:

我已经找到了两篇RAG的论文，我需要再找三篇。

Action: arxiv

Action Input: RAG reading comprehension

获得数据:

Published: 2024-02-19

Title: CRUD-RAG: A Comprehensive Chinese B... ...

正在行动:

我得到的结果是重复的，看来我需要更改一下我的搜索查询以获得不同的结果。

Action: arxiv

Action Input: Retrieval-Augmented Generation

获得数据:

Published: 2012-05-17

Title: On generalized topological groups

A... ...



THANK YOU