

Technical salon / 3rd Anniversary Celebration

CloudWeGo技术沙龙 暨三周年庆典

(3rd _)

Date_

📅 2024/09/21 {Sat} > 14:00_18:00

location

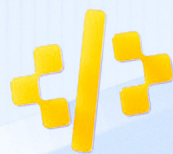
📍 北京大钟寺广场1号楼 F3-19



ABCoder 在大模型编程领域的探索

演讲人：高文举

字节跳动服务框架研发工程师



目录 | Contents

Part 01 我们做了什么

ABCoder Show

Part 03 应用落地

Application(s)

Part 02 ABCoder

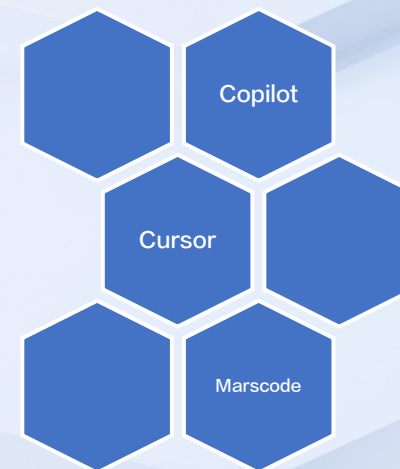
Introduction

Part 04 未来已来?

Open source

Intro

提问：LLM 这把火烧了这么久，大家的编程方式真的改变了吗？



Intro

1. 意图识别不准？

复杂编程问题如何描述和传递？

2. 复杂任务无从下手？

缺乏处理编程任务的标准作业流程（SOP）？

3. 生成效果不理想？

缺失解决问题的完备上下文？

...

...

01

我们做了什么

CloudWeGo x ABCoder Show

我们做了什么

Show case

我们做了什么

小结：

许愿式开发流程：需求沟通 -> IDL 生成 -> 项目生成（Go） -> 项目解析 & 理解
-> 项目部署和测试 -> 项目文档生成

02

ABCoder

初次见面

ABCoder-背景



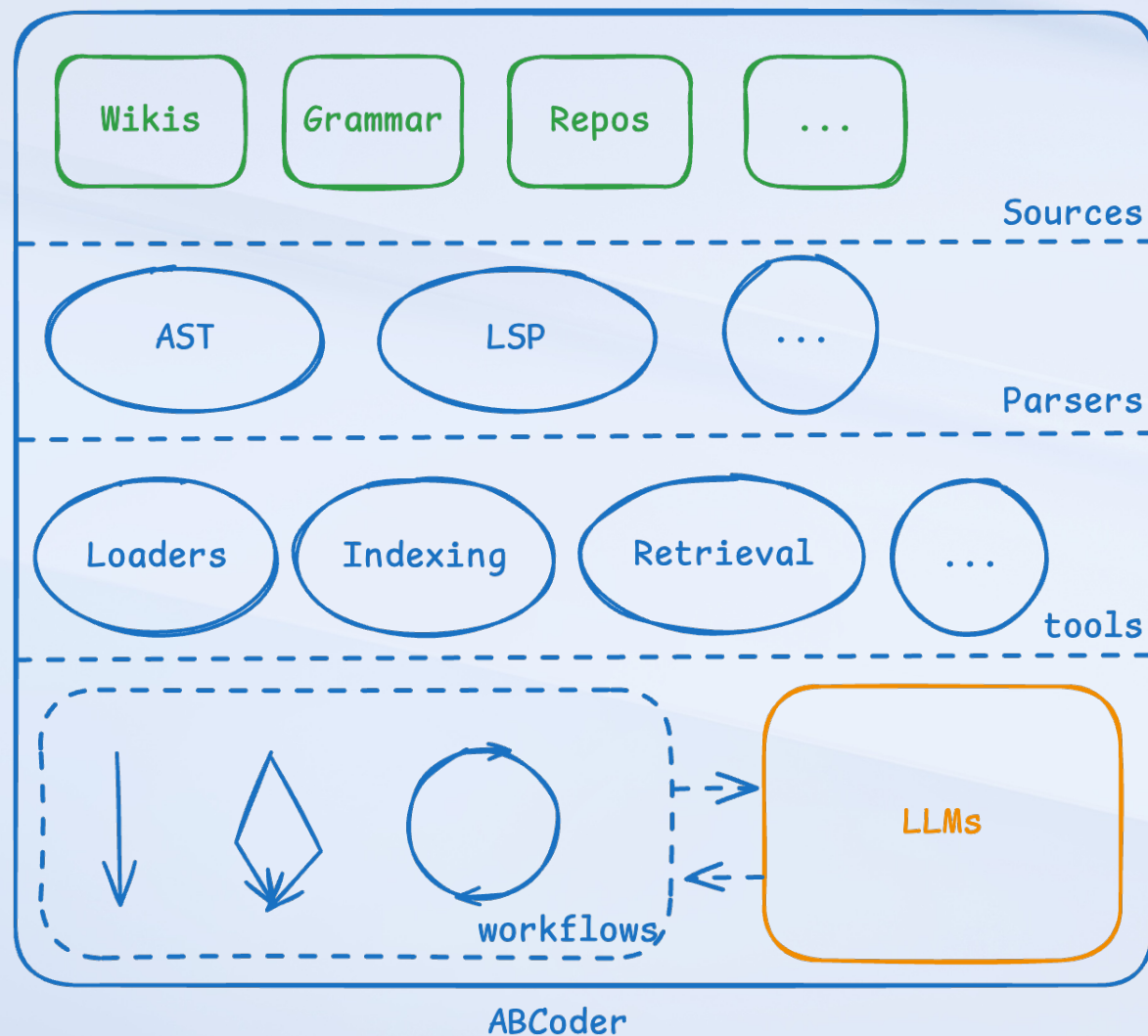
大模型能力涌现：指引了一条充满希望的道路



模型在编程领域的不足：多层次逻辑、复杂算法设计、大型系统架构

ABCoder-内核

ABCoder: AI-Based Coder



编程落地尝试:

- 围绕 LLMs 构建的编程增强解决方案

弥补模型缺乏的编程经验和持续迭代:

- 在复杂逻辑、算法设计及大型系统中稳定发挥其「涌现」



ABCoder-实现

1. 离线路径:

用户文档

编程语言

编程项目



2. 在线路径:

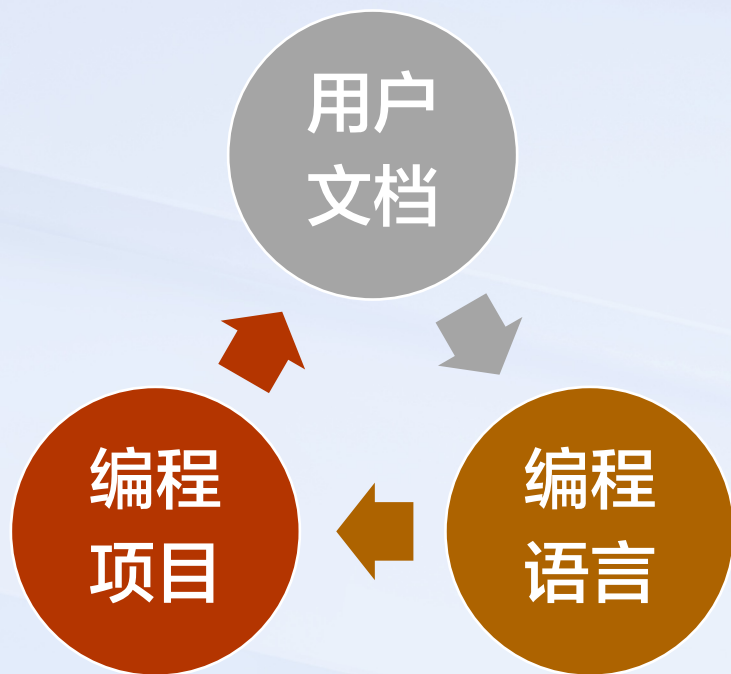


Workflows

工具化



ABCoder-实现（离线）



编程场景 LLM 交互的数据源

1. 用户文档：同主流文档处理流程
2. 编程语言：语法规则 & 工具链
3. 编程项目：ABCoder LLM 原生「解析器」

ABCoder-实现（离线）

离线链路-举个例子

```
.  
├─ README.md  
├─ biz  
│   └─ handler  
│       └─ ping.go  
│   └─ router  
│       └─ register.go  
├─ build.sh  
├─ go.mod  
├─ go.sum  
├─ main.go  
├─ router.go  
├─ router_gen.go  
└─ script  
    └─ bootstrap.sh
```

解析

结构化

压缩

```
main  
├─  
├─ register  
│   └─ router.GeneratedRegister  
│       └─ customizedRegister  
│           └─ handler.Ping  
├─  
└─ server.Default # thirdparty deps
```

Function 调用树/森林

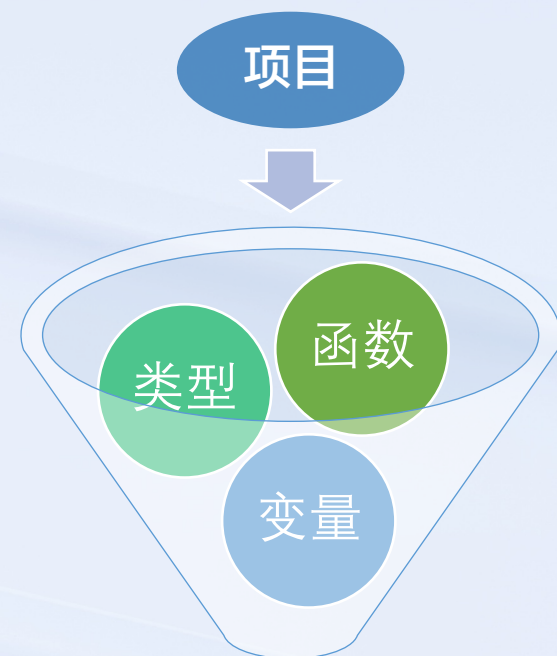
一个典型 Hertz 项目的 Layout

ABCoder-实现（离线）

离线链路-编程项目解析和理解

```
main
|
├─ register
|   ├─ router.GeneratedRegister
|   └─ customizedRegister
|       └─ handler.Ping
|
└─ server.Default # thirdparty deps
```

1. **handler.Ping**: “这是一个处理/ping路由的函数。它返回一个pong的 JSON 响应，表示服务正常。”
2. **customizedRegister**: “该函数用于注册自定义的路由。在目前的实现中，它注册了一个/ping路由，处理函数为handler.Ping。”
3. **router.GeneratedRegister**: “这是一个占位函数，通常用于将生成的路由注册到 Hertz 服务器。目前它没有实际的代码实现，通常是由生成工具填充代码。”
4. **register**: “该函数负责注册所有的路由。它首先调用 router.GeneratedRegister(r)注册生成的路由，然后调用customizedRegister(r)注册自定义路由。”
5. **main**: “该函数是应用程序的入口点。它创建一个默认的 Hertz 服务器实例，然后调用register函数注册路由，最后启动服务器并开始监听请求。”



语言无关的 Schema 抽象

Source code as Knowledge

ABCoder-实现（在线）

Source code
as
Knowledge

- 完备的源码 + LLM 亲和的结构化、语义化知识



Retrieval
&
Augment

- 统一的知识库检索接口设计和实现



Tools
&
Workflows

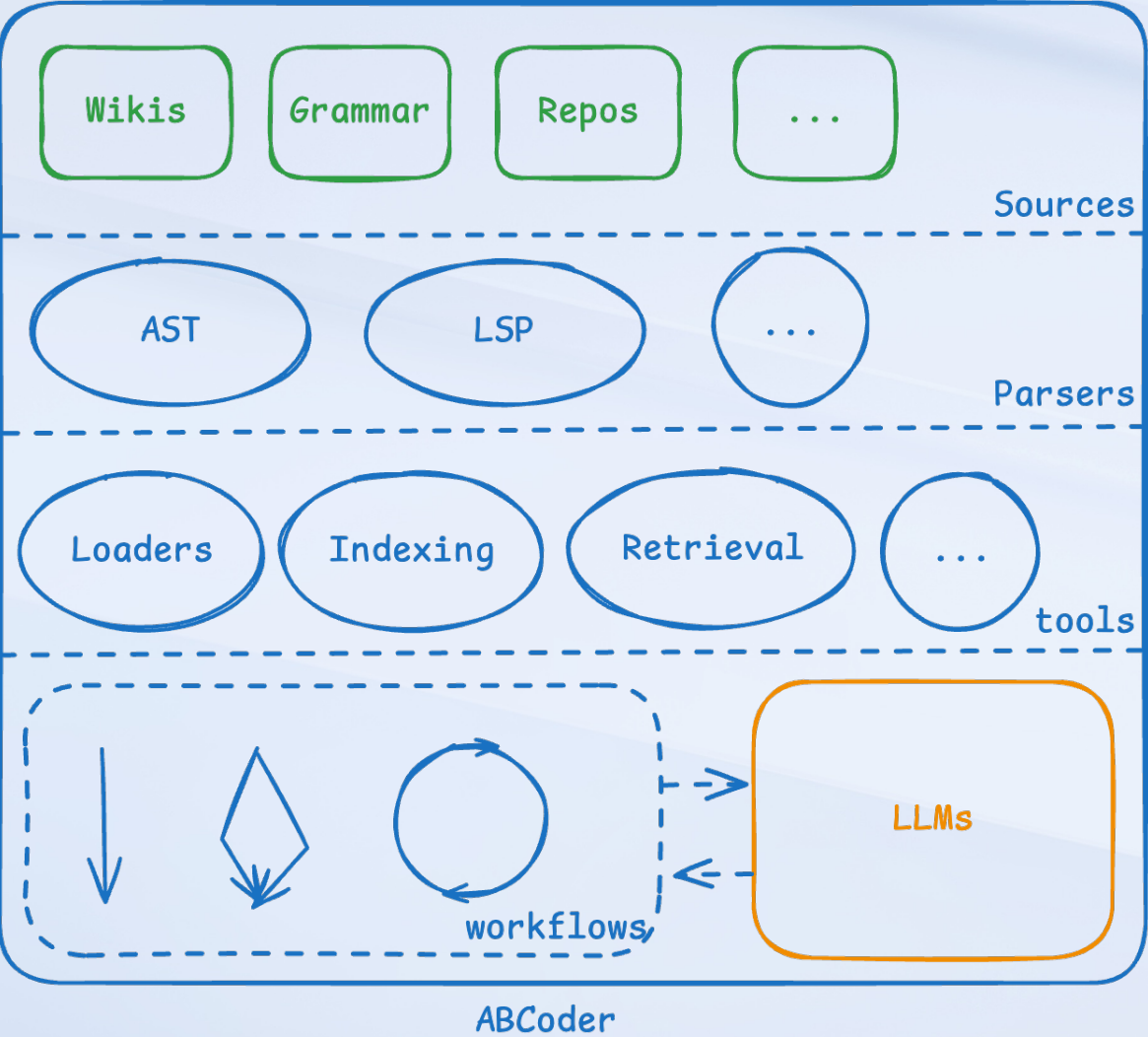
- 工具化 & Workflows 构建和调用

03

应用落地

Application(s)

应用落地-ABCCoder 应用架构



知识库: 食材

Repo parsers: 刀具

Tools: 锅铲

Workflows: 菜谱

LLMs: 厨子

一锅好菜

应用落地-泛化能力

Tools、Workflows 和知识库的组合

应用名称	描述	核心 Tools & Workflows	知识库依赖
Smart Comment	ABCoder 能力 最直接的应用落地 ：项目高质量注释和文档生成	1. ABCoder Repo Loaders 2. 文件系统 Tools	1. ABCoder 解析 & 级联压缩后仓库语料
Wiseman	ABCoder 综合应用 ：具备 CloudWeGo 项目源码级知识和运用能力	同上，额外： 1. ABCoder Parsers 2. IDL tools 3. CMD tools	同上，额外： 1. Examples as Knowledge 2. CloudWeGo 用户文档
A2B（语言翻译）	多系统协作联动应用 ：在 ABCoder 能力之上构建业务导向能力	同上，额外： 1. 复合子系统（IDE 协同、人机协同等）	同上，额外： 1. 更丰富的 语言映射知识库
...			

应用落地-里程碑

Smart comment

近期：产出的注释信息能够达到近乎或者高于人编写的注释质量。

远期：文档工程探索：Code2Wiki

Wiseman

近期：高效解决 CloudWeGo 框架问题的机制，高水准的自闭环率

远期：从需求沟通到自动化研发流

语言翻译（A2B）

近期：持续构建完善的多语言知识库映射，完成从直译到意译的升级

远期：实现渐进式翻译流，完成项目和研发人力 A2B 转换，共同成长

04

未来已来？

开源 & 展望

未来已来? -开源

核心目标

围绕 CloudWeGo 构建 AI 驱动的微服务生态体系

关键路径

1. CloudWeGo 所有组件 ABCoder 索引化
2. ABCoder 应用们走进 CloudWeGo: 研发行为、社区生态 AI 化
3. ABCoder 在 CloudWeGo 完成开源, 拥抱社区

未来已来？—展望

内：ABCoder 的终极理想形态「空」：语言中立编程范式

外：LLM 技术在广度和深度上持续突破：OpenAI O1 = AGI 起点？

THANKS

