

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(10) 国际公布号
WO 2025/123711 A1

(43) 国际公布日
2025 年 6 月 19 日 (19.06.2025)

(51) 国际专利分类号:
G06F 8/30 (2018.01)

(21) 国际申请号: PCT/CN2024/109840

(22) 国际申请日: 2024 年 8 月 5 日 (05.08.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202311723682.4 2023 年 12 月 14 日 (14.12.2023) CN
202410305635.6 2024 年 3 月 15 日 (15.03.2024) CN

(71) 申请人: 华为云计算技术有限公司 (HUAWEI CLOUD COMPUTING TECHNOLOGIES CO., LTD.)
[CN/CN]; 中国贵州省贵阳市贵安新区黔中大道交兴功路华为云数据中心, 550025 (CN)。

(72) 发明人: 申博 (SHEN, Bo); 中国贵州省贵阳市贵安新区黔中大道交兴功路华为云数据中心 550025 (CN)。郭悦楠 (GUO, Yuenan); 中国贵州省贵阳市贵安新区黔中大道交兴功路华为云数据中心 550025 (CN)。俞蔼伦 (YU, Ailun); 中国贵州省贵阳市贵安新区黔中大道交兴功路华为云数据中心 550025 (CN)。

(74) 代理人: 深圳市深佳知识产权代理事务所 (普通合伙) (SHENPAT INTELLECTUAL PROPERTY AGENCY); 中国广东省深圳市罗湖区南湖街道春风路庐山大厦 B 座 18C2、18D、18E、18E2 518001 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI,

(54) Title: CODE GENERATION METHOD AND RELATED DEVICE

(54) 发明名称: 一种代码生成方法及相关设备

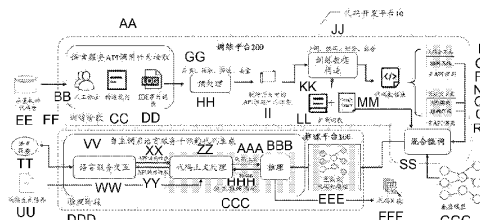


图 1

BBB Inference
CCC Streaming inference loop
DDD Inference stage
EEE Generative code large model
FFF Code snippets
GGG Base model
HHH token

10 Code development platform
100 Inference platform
200 Training platform
AA Language service API calling behavior extraction
BB Manual annotation
CC Rule construction
DD IDE event recording
EE Massive software code repository
FF Training stage
GG Classification, extraction, screening and de-duplication
HH Pre-processing
II API calling behavior information in software development
JJ Word segmentation, encoding, segmentation and combination
KK Training data construction
LL Word list expansion
MM Code data set
NN, AAA Code context
OO Code prediction
PP Have no API calling
QQ API calling
RR Have API calling
SS Hybrid fine-tuning
TT Language service
UU Code generation task
VV Code generation intervened by autonomously calling the language service
WW Language service interaction
XX API calling response
YY API calling request
ZZ Code context processing

(57) Abstract: Disclosed in the present application is a code generation method. The method comprises: a code development platform receiving a code generation task, and extracting code context on the basis of the code generation task, wherein the code context comprises code snippets in a project which are related to a code to be generated; the code development platform determining, on the basis of the code context and by means of a language model, whether to call an application programming interface (API) of a language service; when determining to call the API of the language service, the code development platform sending an API calling request to the language service, so as to obtain an API calling response from the language service; and on the basis of the code context and the API calling response from the language service, the code development platform performing inference by means of the language model, so as to obtain a generated code. The method endows a language model with the capability of autonomously calling an external tool, e.g., a language service, such that the language model can simulate, during inference, a human developer to use the language service to assist

[见续页]

GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

the language model in code generation, thereby alleviating and even eliminating the problem of hallucination caused by the language model during code generation.

(57) 摘要: 本申请公开一种代码生成方法, 包括: 代码开发平台接收代码生成任务, 根据代码生成任务提取代码上下文, 代码上下文包括项目中与待生成代码相关的代码片段, 代码开发平台根据代码上下文, 通过语言模型确定是否调用语言服务的应用程序编程接口API。当确定调用语言服务的API, 代码开发平台向语言服务发送API调用请求, 获得语言服务的API调用响应, 代码开发平台根据代码上下文和语言服务的API调用响应, 通过语言模型推理以获得生成代码。该方法赋予语言模型自主调用语言服务等外部工具的能力, 使得语言模型在推理的过程中能够模拟人类开发者使用语言服务辅助该语言模型进行代码生成, 缓解甚至杜绝语言模型在代码生成过程中产生的幻觉问题。