前　言

近年来,大语言模型(LLM)取得了突飞猛进的发展,将人工智能推向了一个新的高度,使更多不同的应用和工业领域获得了智能,并增强了能力。人机交互更加自然高效,各行业出现新的服务和业务模式,科学研究也得到加速。大模型背后是超大的参数规模、海量的训练数据和强大的计算资源,但软件能力提升的同时其质量保障也越来越困难。相比传统软件工程,大模型时代的软件工程需要面临更多的挑战:大模型支撑系统与框架的潜在缺陷,不断变化的基础设施导致系统演化困难,大模型行为本身缺乏可解释性与可信性,数据质量问题诱发的模型质量问题,伦理、道德和法律的约束使得软件应用困难。在这种局面下,亟需从软件工程的角度系统性地审视大模型时代下的软件质量保障方法,从而更好应对这一新兴领域中的新挑战。“大模型与软件工程”专题立足于软件工程领域,探讨在大模型时代软件工程技术的新发展与新趋势,特别是基于大模型的需求分析、设计和如何利用大模型技术进行自动代码生成、测试和形式化验证等,同时研究如何利用软件工程理论和方法保证基于大模型的软件系统的质量和安全。

本专题公开征文,共收到投稿51篇。论文均通过了形式审查,内容涉及大语言模型赋能的软件需求分析、设计、编码、测试与维护等软件全生命周期。特约编辑先后邀请了60多位专家参与审稿工作,每篇稿件至少邀请3位专家进行评审。稿件经初审、复审、ChinaSoft 2025会议宣读和终审4个阶段,历时6个月,最终有以下6篇论文入选本专题。

高俊涛等作者的“基于大语言模型的需求歧义检测方法”针对现有需求歧义检测方法存在歧义检测解决方案类型不全,难以全面覆盖语言歧义对应的6种歧义类型,以及传统方法缺乏深层次的语义理解、逻辑关系识别以及引用/指代关系识别能力等问题,提出了一种基于大语言模型的方法LMAdetect,用于自动检测需求歧义。对于给定的需求,LMAdetect根据启发式规则和大语言模型进行分析,将该需求检测分类为对应的歧义类型。针对大语言模型的分类结果,LMAdetect利用基于规则的算法,将不同歧义类型分配给不同专家进行检测。最后对这些检测结果根据置信度进行汇总并输出置信度高的分类结果。

谭舒孺等作者的“RD2ESC:多智能体嵌入式代码生成框架”是针对嵌入式系统特殊需求的有效解决方案。提出了RD2ESC(requirements documents to embedded system code)方法,通过基于提示词的微调技术使大语言模型能够理解嵌入式代码与需求文档之间的复杂关系,并构建了多智能体协同的代码生成框架,能够利用需求文档和参考代码快速生成高质量的嵌入式代码。该研究为嵌入式代码自动生成提供了有效的技术框架,提升了嵌入式系统开发效率。

嵇友晴等作者的“融合静态分析与大语言模型的非连续代码重构方法”针对于由语句交错、重排等导致的非连续代码克隆,大语言模型在重构时面临着语义上下文分散、关键依赖捕捉困难以及易产生“幻觉”错误等核心挑战,提出了一种融合静态分析与大语言模型的非连续代码克隆重构方法。该方法首先通过结合程序切片与代数分类器,高效精准地识别非连续克隆;然后,通过一种基于上下文信息的重构机会识别算法,为大语言模型确定最佳重构目标;最后,利用思维链少样本提示策略引导大语言模型生成高质量的“提取函数”重构建议,并利

用蜕变关系验证机制对生成结果进行语义和结构一致性验证。

张瑞宸等作者的“基于大语言模型的数据库管理系统模糊测试方法”针对数据库管理系统（database management system, DBMS）现有的DBMS模糊测试方法面临测试用例覆盖能力不足与多DBMS适配性差两大问题，提出CLCC（curated LLM case construct），一种基于大语言模型的DBMS模糊测试方法。该方法在模糊测试前，利用大语言模型对初始种子进行构建，并在模糊测试过程中，根据边覆盖情况筛选种子引导大语言模型生成测试用例。

魏威等作者的“多代理协作实现缺陷报告补全和优化”受人类专家“分阶段处理、多角色协同”的问题解决机制启发，提出一种多代理协作机制实现缺陷报告补全和优化。通过3个关键设计实现高质量补全：1）将补全任务分解为缺陷分析、报告补全和质量评估3个阶段，由不同代理负责以降低认知负荷；2）设计结构化提示模板，精准引导大语言模型扮演领域专家角色，确保各阶段输出准确；3）引入动态反馈机制，通过多轮迭代实现代理间的交叉验证和协同优化，有效控制语义漂移问题，确保逻辑连贯且补全内容与事实一致。

崔星等作者的“基于代码感知与双阶段优化融合的README生成大语言模型框架”提出结合大语言模型与代码结构建模的双阶段README自动生成框架RMancer。第1个阶段中，RMancer设计了基于提示引导的结构化信息抽取方法，结合静态分析生成高质量训练数据，提升模型对文件级功能摘要、依赖关系和主程序入口等结构要素的感知能力。第2个阶段设计了基于调用图的拓扑排序策略，重构模块间的执行逻辑顺序，以构建结构化文档生成的上下文信息；同时，引入多任务监督机制，引导大模型联合学习文档段落结构与内容生成，提升输出文本的逻辑一致性与客观性；最后，RMancer通过标准化约束策略对生成结果进行格式规整与内容审校，确保文档的规范性与准确性。

本专题主要面向软件工程、系统软件、形式化、人工智能和大模型等多领域的研究人员和工程人员，反映了我国学者在大语言模型、人工智能和软件工程等领域相互赋能最新的研究进展。感谢《计算机研究与发展》编委会和软件工程专委会对专题工作的指导和帮助，感谢专题全体评审专家及时、耐心、细致的评审工作，感谢踊跃投稿的所有作者。希望本专题能够对与大语言模型相互赋能的软件工程、系统软件与形式化等相关领域的研究工作有所促进。

聂长海（南京大学）
王璐（西安电子科技大学）
王莹（东北大学）
姜艳杰（天津大学）
张敏灵（东南大学）
2026年3月