Structured Knowledge meets GenAI: A Framework for Logic-Driven Language Models

Farida Eldessouky[1], Nourhan Ehab[1], Carolin Schindler[2], Mervat Abuelkheir[1], Wolfgang Minker[2]
[1]German University in Cairo, [2]Ulm University
Correspondence: farida.el-dessouky@guc.edu.eg

Abstract

Large Language Models (LLMs) excel at generating fluent text but struggle with context sensitivity, logical reasoning, and personalization without extensive fine-tuning. This paper presents a logical modulator: an adaptable communication layer between Knowledge Graphs (KGs) and LLMs as a way to address these limitations. Unlike direct KG-LLM integrations, our modulator is domain-agnostic and incorporates logical dependencies and commonsense reasoning to achieve contextual personalization. By enhancing KG interaction, this method will produce linguistically coherent and logically sound outputs, increasing interpretability and reliability in generative AI.

1 Introduction

As LLMs gain prominence in generating natural language, surveys point out that their reasoning capabilities become increasingly apparent (Chang et al., 2024). When dealing with critical applications like healthcare (Nazi and Peng, 2024), trust is vital and well-informed decisions need to be guaranteed. Hence, the gaps related to the black box nature of LLMs highlight the need for models that cannot only generate human-like text but also make informed, context-aware decisions. Unlike existing methods that attempt to incorporate KGs as external sources (Sui and Hooi, 2024), our approach centers on a bidirectional mediator that enables a dynamic and adaptable exchange between the LLM and KG. The KG serves as the primary source for grounded, factual, and explainable information, while the LLM provides the necessary fluency and cohesiveness to translate structured data into human-readable language. This framework ensures that responses are not only accurate but also explainable, aligning well with applications where trust and transparency are essential.

2 Related Work

LLMs rely on statistical patterns learned through extensive amounts of data, rather than true logical reasoning. This can lead to errors in context-sensitive tasks (Zhou et al., 2024). This limitation is especially problematic in sensitive fields like medicine and law (Wang et al., 2023), where accuracy and logical consistency are crucial. Approaches following retrieval-augmented generation (RAG) (Lewis et al., 2020) revolutionize factual accuracy by incorporating external knowledge. However, they primarily rely on surface-level matching and embedding similarities, lacking the depth needed for complex reasoning (Mao et al., 2021). Graph-based enhancements of this approach, such as GraphRAG (Peng et al., 2024), offer structured knowledge integration through KGs but reduce KG data to vector embeddings, which strips away important logical dependencies. Some studies explore utilizing KGs as a prompt mechanism to support graph-based reasoning tasks in LLMs. (Zhang, 2023; Huang et al., 2024). They find that embedding graph data into LLM prompts can improve reasoning capabilities, while there is also a key limitation: The integration often fails to retain the full relational structure of the KG, leading to limitations in multi-step reasoning and contextual consistency.

3 A Framework for KG-Enhanced LLM Reasoning

To enhance reasoning and personalization in LLMs, we propose a domain-independent end-to-end framework centered around an independent mediator as shown in Figure 1. This mediator can be perceived as a way to perform the retrieval part of RAG. The main difference in our architecture from traditional methods is that it better structures the reasoning behind the retrieval through a decompositional querying mechanism. This way a more optimal use of structured knowledge can be

Proceedings of the Generative AI and Knowledge Graph Workshop (GenAIK), pages 66–68
January 19, 2025. ©2025 International Committee on Computational Linguistics (ICCL)

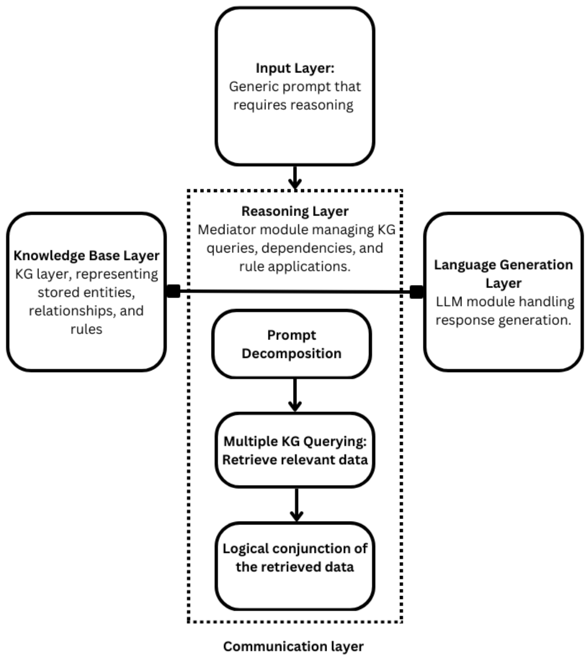

Figure 1: Modular Architecture for our proposed end-end framework for KG-enhanced LLM reasoning.

achieved. The mediator breaks down user inputs into KG queries, interprets KG responses for the LLM by extracting the needed nodes and relationships (which can be not very prose friendly), and handles the back and forth communication between separate LLM and KG modules to achieve comprehensive answers. By coordinating these interactions, the mediator enables the KG to act as the primary source of reasoning, while the LLM provides fluent language generation. This design offers a versatile solution that addresses gaps in LLM reasoning and adapts seamlessly to various domains. Separating the modules in our systems and connecting the dots through our proposed modulator organizes the communication between these modules, making the pipeline more flexible and interpretable. With this approach, we aim at aligning outputs with user-specific data from the KG, preserving logical depth that simpler KG-LLM integrations lack. Moreover, integrating symbolic AI with LLMs supports explainability, a key issue in AI (Longo et al., 2024), by allowing responses to be traced back to specific KG elements and rules, enhancing transparency and fostering trust. Furthermore, we aim for organizing the pipeline by separating the modules and facilitating the communication between them. Unlike retrieval-based approaches (e.g., GraphRAG, LightRAG (Guo et al., 2024)), which reduce KG data to vector representations, our framework directly interacts with the KGs through a dynamic reasoning layer. This hybrid approach combines the structured knowledge of KGs with the linguistic capabilities of LLMs, improving factual accuracy, logical depth, and personalization.

4 Conclusion and Future Work

By introducing a decompositional reasoning layer to interface with KGs, this research offers a novel approach to improve reasoning in LLMs. Our method targets responses that are contextually grounded and linguistically fluent by giving LLMs fine-grained access to structures of KGs. Applications where sophisticated reasoning is crucial are particularly promising for this hybrid architecture. KG-enhanced models that adjust to particular user demands while retaining logical accuracy could be extremely helpful in domains including career development, healthcare planning, and legal advising. These domains offer use-cases to be explored when conducting empirical studies with our approach. Moreover, since knowledge needs to be updated regularly in practical applications, future work includes investigating how KG completion with LLMs can benefit from our proposed reasoning mediator. This way, the bidirectional property of the mediator is leveraged to dynamically enrich the KG with new, contextually relevant knowledge, without the need to manually modify it.

Limitations and Ethical Considerations

Integrating KGs with LLMs, despite its advantages, faces key challenges. Real-time KG querying can introduce latency, especially when responses require multiple reasoning steps or extensive entity retrieval. Accurate entity linking is also complex due to ambiguity, synonyms, and domain-specific terms, requiring advanced learning for reliable mapping, as misalignment can reduce response relevance. Scalability poses further issues, as expanding the KG for broader knowledge increases storage and rule complexity. Additionally, evaluating reasoning quality, particularly for personalized tasks, often demands expert review, limiting scalability. Overcoming these obstacles is crucial for a robust, high-performance KG-augmented LLM system. Ensuring transparency in KG construction, regular bias mitigation, and accountability in how responses are generated will be essential to address ethical concerns.

References

Yupeng Chang, Xu Wang, Jindong Wang, Yuan Wu, Linyi Yang, Kaijie Zhu, Hao Chen, Xiaoyuan Yi, Cunxiang Wang, Yidong Wang, et al. 2024. A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology*, 15(3):1–45.

Zirui Guo, Lianghao Xia, Yanhua Yu, Tu Ao, and Chao Huang. 2024. Lightrag: Simple and fast retrieval-augmented generation. *arXiv preprint arXiv:2410.05779*.

Wenyu Huang, Guancheng Zhou, Mirella Lapata, Pavlos Vougiouklis, Sebastien Montella, and Jeff Z Pan. 2024. Prompting large language models with knowledge graphs for question answering involving long-tail facts. *arXiv preprint arXiv:2405.06524*.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, 33:9459–9474.

Luca Longo, Mario Brcic, Federico Cabitza, Jaesik Choi, Roberto Confalonieri, Javier Del Ser, Riccardo Guidotti, Yoichi Hayashi, Francisco Herrera, Andreas Holzinger, et al. 2024. Explainable artificial intelligence (xai) 2.0: A manifesto of open challenges and interdisciplinary research directions. *Information Fusion*, 106:102301.

Yuning Mao, Pengcheng He, Xiaodong Liu, Yelong Shen, Jianfeng Gao, Jiawei Han, and Weizhu Chen. 2021. Generation-augmented retrieval for open-domain question answering. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 4089–4100, Online. Association for Computational Linguistics.

Zabir Al Nazi and Wei Peng. 2024. Large language models in healthcare and medical domain: A review. In *Informatics*, volume 11, page 57. MDPI.

Boci Peng, Yun Zhu, Yongchao Liu, Xiaohe Bo, Haizhou Shi, Chuntao Hong, Yan Zhang, and Siliang Tang. 2024. Graph retrieval-augmented generation: A survey. *arXiv preprint arXiv:2408.08921*.

Yuan Sui and Bryan Hooi. 2024. Can knowledge graphs make large language models more trustworthy? an empirical study over open-ended question answering. *arXiv preprint arXiv:2410.08085*.

Cunxiang Wang, Xiaoze Liu, Yuanhao Yue, Xiangru Tang, Tianhang Zhang, Cheng Jiayang, Yunzhi Yao, Wenyang Gao, Xuming Hu, Zehan Qi, et al. 2023. Survey on factuality in large language models: Knowledge, retrieval and domain-specificity. *arXiv preprint arXiv:2310.07521*.

Jiawei Zhang. 2023. Graph-toolformer: To empower llms with graph reasoning ability via prompt augmented by chatgpt. *arXiv preprint arXiv:2304.11116*.

Bo Zhou, Daniel Geißler, and Paul Lukowicz. 2024. Misinforming llms: vulnerabilities, challenges and opportunities. *arXiv preprint arXiv:2408.01168*.