LLM-Powered Graph Reasoning for Knowledge Discovery

Knowledge graphs (KGs) encode structured relations between entities and are widely used for scientific discovery. Graph Neural Networks (GNNs) can effectively encode KGs for link prediction and relational modeling, but they are computationally expensive and do not generalize well to unobserved entities. Large language models (LLMs) excel at zero-shot reasoning and generalization to unseen data, but cannot directly process large KGs. Together, these limitations highlight key challenges in using large KGs for prediction tasks: (1) Scalability–their growing scale and continuous evolution with new entities make them computationally expensive to process and train on. (2) Generalization–most existing methods fail to handle previously unseen entities. These challenges motivate a central research problem: **How can we extract compact, interpretable reasoning paths from large KGs so they can be effectively used by both GNNs and LLMs for prediction tasks?**

To address these problems, we introduce K-Paths (Figure 1.1), a training-free framework that retrieves diverse multi-hop reasoning paths between entities. It prunes the KG for task-specific interactions by applying a diversity-aware Yen's algorithm [1] to extract non-redundant paths. These paths are provided as subgraphs for GNNs or natural language descriptions for LLMs, reducing computation, enabling generalization to new entities, and offering interpretable reasoning evidence.

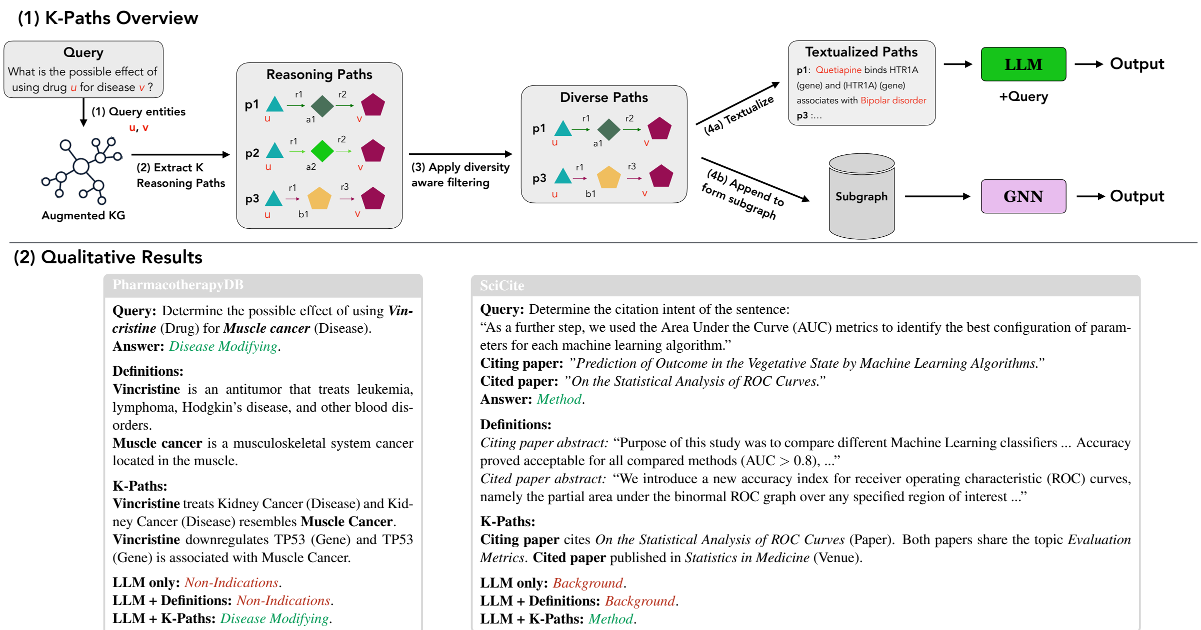

Figure 1: Overview of K-Paths and qualitative results. (1) K-Paths extracts diverse reasoning paths between entities u and v, applies diversity filtering, and outputs them as text for LLMs or subgraphs for GNNs. (2) Examples across datasets show that these interpretable paths guide LLMs to correct predictions.

We applied K-Paths to diverse biomedical and academic KGs. On drug–based prediction tasks, K-Paths yields large zero-shot gains for LLMs: Llama 3.1 8B improves from 14.7 to 47.0 F1-score on *DDInter* and from near-zero to 40.5 F1-score on *DrugBank*, while Tx-Gemma models gain +25–37 F1-score. For supervised GNNs, K-Paths cuts graph size by $\sim$90% while preserving or improving accuracy, e.g., EmerGNN [2] improves from 68.0 to 68.9 F1-score on *DDInter*, while training time drops by 83%. On citation intent classification tasks, K-Paths improves LLM F1-score by 6–13 points on *SciCite*, demonstrating cross-domain generalization. Qualitative results (Figure 1.2) illustrate how K-Paths grounds predictions with reasoning evidence. In summary, K-Paths reduces the computational cost of large KGs, improves generalization to unseen entities, and provides human-readable reasoning paths for explainability. Future work includes better path scoring, joint retrieval–reasoning, and broader applications in scientific discovery.

[1] J. Y. Yen. "Finding the K Shortest Loopless Paths in a Network", 1971.

[2] Y. Zhang et al. "Emerging Drug Interaction Prediction Enabled by Flow-based Graph Neural Network with Biomedical Network", 2023.