Theory and Practice of Science and Technology
2026, VOL. 7, NO. 2, 132-135
DOI: 10.47297/taposatWSP2633-456926.20260702

Knowledge Graph-Empowered Natural Language Inference: Methods, Applications, and Frontiers

Wei Li

China University of Political Science and Law, Haidian District, Beijing, 100088, China

ABSTRACT
Natural Language Inference (NLI), also known as Recognizing Textual Entailment, is a fundamental task in natural language processing that involves determining the relationship between two text segments: a premise and a hypothesis. The goal is to classify whether the premise entails, contradicts, or is neutral with respect to the hypothesis. While traditional NLI models primarily rely on the textual information within the premise and hypothesis, the integration of external knowledge, particularly in the form of knowledge graphs (KGs), has emerged as a promising direction. Knowledge graphs provide structured representations of real-world facts and relationships, enabling NLI models to reason beyond the surface-level semantics of the text. This review paper offers a detailed examination of knowledge graph-empowered NLI, encompassing various methodologies, applications, and future research directions. We begin by providing a historical overview of NLI and the evolving role of knowledge integration. We then delve into the core methods, categorizing them based on KG embedding techniques and KG-augmented attention mechanisms. We explore diverse applications of KG-enhanced NLI across domains like question answering, dialog systems, and misinformation detection. Furthermore, we address the challenges associated with KG-NLI, including knowledge graph incompleteness and scalability issues. Finally, we discuss promising future perspectives, such as incorporating commonsense knowledge and developing more robust KG reasoning techniques. This comprehensive review aims to provide researchers and practitioners with a thorough understanding of the current state and future potential of knowledge graph-empowered natural language inference.

KEYWORDS
Natural language inference; Knowledge graph; NLI, Knowledge representation; Textual entailment; Reasoning; Deep learning

1 Introduction

1.1 Background and Motivation

Natural Language Inference (NLI) is a fundamental task in natural language processing, aiming to determine the semantic relationship between two sentences: a premise and a hypothesis [1]. Accurate NLI models are crucial for various downstream applications, including question answering, text summarization, and dialogue systems [2]. However, traditional NLI models often struggle with inferences requiring external knowledge beyond the text. Incorporating knowledge graphs (KGs) offers a promising avenue to enhance NLI by providing structured background knowledge. KGs can help models reason about entities and relations, leading to improved accuracy and robustness, especially when dealing with complex or implicit inferences where $p \Longrightarrow$h.

1.2 Knowledge Graphs for NLI: An Overview

Knowledge Graphs (KGs) represent structured knowledge as entities and relations, offering a powerful resource for enhancing Natural Language Inference (NLI). By encoding factual information and semantic relationships, KGs provide external knowledge to enrich the understanding of text. Different types of KGs exist, including general-purpose KGs like Wikidata and domain-specific KGs tailored to particular areas. KGs are characterized by their scale (N entities,M relations), density (average connectivity), and schema (types of entities and relations), influencing their utility in NLI tasks.

2 Historical Overview of NLI and Knowledge Integration

2.1 Early Approaches to NLI

Early NLI systems relied on rule-based methods, employing hand-crafted rules to determine entailment. These systems often used lexical matching and simple logical inferences. Statistical approaches, such as those based on bag-of-words models and n-grams, emerged later, learning patterns from data. However, both approaches exhibited limitations [3]. Rule-based systems struggled with scalability and coverage, while statistical methods lacked the ability to reason about complex relationships and contextual nuances. The inability to capture world knowledge and common-sense reasoning highlighted the need for integrating external knowledge sources to improve NLI accuracy and robustness (Table 1).

Table 1 Evolution of NLI Datasets

Approach	Strengths	Weaknesses	Need
Rule-based Methods	Simple logical inferences, lexical matching	Scalability, coverage, hard to maintain	External knowledge sources for improved accuracy and reasoning.
Statistical Methods	Learning patterns from data, automated	Inability to reason about complex relationships, lack of contextual nuances	Integration of world knowledge and common-sense reasoning.

2.2 The Rise of Deep Learning and Knowledge-Aware Models

The advent of deep learning revolutionized NLI, moving away from feature engineering towards end-to-end trainable models [4-5]. Recurrent Neural Networks (RNNs), particularly LSTMs, became foundational, encoding sentences into vector representations. Attention mechanisms further enhanced performance by allowing models to focus on relevant parts of the input. Initial attempts to incorporate external knowledge into these deep learning architectures often involved embedding knowledge graph entities and relations into vector spaces, using techniques like TransE. These embeddings were then integrated into the NLI model, typically by concatenating them with sentence embeddings or using them to guide attention mechanisms. The goal was to provide the model with additional semantic information beyond what was present in the text, enabling better reasoning and inference capabilities. However, effectively fusing symbolic knowledge with distributed representations remained a significant challenge [6].

3 KG Embedding-Based NLI Methods

3.1 Knowledge Graph Embeddings for NLI

Knowledge graph embeddings offer a powerful approach to incorporating structured knowledge into Natural Language Inference (NLI) models [7]. These techniques represent entities and relations within a knowledge graph as low-dimensional vectors, capturing semantic information crucial for inference [8]. Methods like TransE learn embeddings by treating relations as translations between entity embeddings; ideally, if (h,*r*,*t*) exists in the KG, then $\mathrm{h} + r \approx t$ in the embedding space. ComplEx extends this by using complex-valued embeddings to handle asymmetric relations. RotatE further enhances relation modeling by interpreting relations as rotations in a complex space, where $t = \mathrm{h} \circ r$ (element-wise multiplication). In NLI, these embeddings can be used to represent entities and relations extracted from both the premise and the hypothesis [9]. By comparing the embeddings of corresponding entities and relations, the model can assess the semantic relatedness between the two sentences. For example, if the premise contains the fact "apple isA fruit" and the hypothesis states "apple is a type of fruit", the embeddings of "apple", "isA", and "fruit" can be used to determine the compatibility between the premise and hypothesis, leading to a more accurate inference. The learned embeddings thus provide a structured semantic representation that complements the textual information, improving NLI performance (Table 2).

Table 2 Comparison of KG Embedding Techniques

Technique	Relation Modeling	Embedding Space	Key Idea	Strengths	Weaknesses
TransE	Translation	Real-valued	$\mathrm{h} + r \approx t$	Simple and efficient.	Struggles with complex relations (e.g., one-to-many).
ComplEx	Bilinear product	Complex-valued	Handles asymmetric relations through complex conjugations.	Improved handling of asymmetric relations compared to TransE.	Can be computationally expensive.
RotatE	Rotation	Complex-valued	$t = \mathrm{h} \circ r$ (element-wise multiplication)	Models various relation patterns (symmetry/antisymmetry, inversion, composition).	More complex to train; potential for overfitting.

3.2 Integrating KG Embeddings into NLI Architectures

Integrating knowledge graph embeddings into NLI architectures involves injecting entity and relation information extracted from KGs into various stages of the NLI model. One common approach is to enhance the input representation. For instance, entity embeddings, representing the entities present in the premise and hypothesis, can be concatenated with word embeddings or fed into separate embedding layers. This allows the model to directly access entity-level semantics. Another strategy focuses on incorporating relation information. Relation embeddings, representing the relationships between entities, can be used to augment the input or hidden layers. For example, relation embeddings can be concatenated with the hidden states of the premise and hypothesis, enabling the model to consider the relationships between entities when reasoning about the entailment relation [10]. Attention mechanisms can also be employed to selectively attend to relevant entities and relations based on the context. The attention weights, a_{ij}, reflect the importance of the *j*-th entity given the *i*-th word. Furthermore, graph neural networks (GNNs) can be used to encode the KG structure and generate contextualized entity and relation embeddings, which are then integrated into the NLI model. These GNN-derived embeddings capture higher-order relational information and improve the model's reasoning capabilities.

3.3 Case Studies and Performance Analysis

KG embedding-based NLI models demonstrate notable gains in scenarios requiring commonsense reasoning. For instance, incorporating ConceptNet embeddings significantly improves performance on datasets like SNLI and MNLI, particularly when dealing with hyponymy and meronymy relations. However, performance often plateaus on examples demanding complex logical inference or nuanced contextual understanding. A trade-off exists between embedding dimensionality and computational cost; higher dimensional embeddings capture richer semantic information but increase training time and memory requirements. Furthermore, these models are susceptible to biases present in the underlying knowledge graphs, potentially leading to unfair or inaccurate predictions. The effectiveness also depends

heavily on the quality and coverage of the KG; incomplete or noisy KGs can hinder performance.

4 KG-Augmented Attention Mechanisms for NLI

4.1 Attention Mechanisms and Knowledge Integration

Attention mechanisms provide a powerful means to selectively focus on the most relevant parts of a knowledge graph (KG) when performing natural language inference (NLI). By attending to specific entities and relations within the KG, these mechanisms enable models to dynamically integrate knowledge relevant to both the premise and the hypothesis. This selective attention is crucial because not all information within the KG is pertinent to a given NLI task; attending to irrelevant information can introduce noise and hinder performance. Specifically, attention weights are typically computed based on the similarity between the premise/hypothesis representations and the embeddings of entities or relations in the KG. These weights then determine the contribution of each KG component to the final representation used for inference. Furthermore, KG-augmented attention mechanisms excel at capturing long-range dependencies and contextual information. The KG provides a structured representation of relationships between entities, allowing the model to infer connections that might not be explicitly stated in the premise or hypothesis. For example, if the premise mentions entity A and the hypothesis mentions entity C, the attention mechanism can leverage the KG to identify a relevant path $A \rightarrow B \rightarrow C$, thereby enriching the contextual understanding and improving inference accuracy. The attention score a_{ij} reflects the importance of the j-th KG component to the i-th word in the premise or hypothesis.

4.2 Graph Attention Networks for NLI

Graph Attention Networks (GATs) offer a powerful mechanism for reasoning over knowledge graphs within Natural Language Inference (NLI) tasks. Their core strength lies in their ability to selectively aggregate information from neighboring entities and relations, assigning different importance weights to each neighbor based on learned attention coefficients. This is particularly beneficial in NLI, where understanding the relationships between entities mentioned in the premise and hypothesis is crucial for determining the inference relationship.

Specifically, GATs compute attention weights a_{ij} between node i and its neighbor j based on their features. These weights, often normalized using a softmax function, represent the importance of neighbor j to node i. The aggregated representation of node i is then a weighted sum of its neighbors' features, allowing the model to focus on the most relevant connections within the knowledge graph. This attention mechanism enables GATs to capture complex relationships and dependencies, leading to improved performance in NLI tasks by providing a more nuanced understanding of the semantic relationships encoded in the knowledge graph. The learned attention weights can also provide insights into which relationships are most important for making inference decisions (Table 3).

Table 3 GAT Layers for Knowledge Aggregation

Component	Description	Benefit in NLI
Attention Weights a_{ij}	Represents the importance of neighbor j to node i, based on their features. Computed and normalized (e.g., using softmax).	Enables selective information aggregation, focusing on relevant connections between entities in the premise and hypothesis.
Weighted Sum of Neighbors' Features	Aggregation process where the representation of node i is calculated using the attention-weighted features of its neighbors.	Captures complex relationships and dependencies within the knowledge graph, leading to a more nuanced understanding of semantic relationships.
Learned Attention Weights	Attention weights that are learned during training, indicating the strength of relationship between two nodes.	Provides insights into which relationships are most important for making inference decisions in NLI tasks.

4.3 Comparative Study of Attention-Based Architectures

Attention mechanisms are crucial for integrating knowledge graph information in NLI. Architectures vary in how they attend to graph entities and relations. Some models employ separate attention layers for premise and hypothesis, while others use a unified attention mechanism. Graph attention networks (GATs) excel at capturing complex relational dependencies but can be computationally expensive. Simpler attention mechanisms, like scaled dot-product attention, offer efficiency but may sacrifice accuracy on complex inferences. Performance often depends on the size and quality of the knowledge graph and the complexity of the NLI task. The choice of architecture involves a trade-off between model complexity, computational cost, and performance gains on specific datasets.

5 Comparison, Challenges, and Limitations

5.1 Comparative Analysis of KG-NLI Approaches

KG embedding-based NLI models typically encode entities and relations into low-dimensional vector spaces, enabling efficient reasoning through vector operations. These methods excel in capturing semantic relatedness and handling large-scale KGs, but often struggle with complex logical inferences requiring multi-hop reasoning. Conversely, KG-augmented attention mechanisms directly incorporate KG information into neural network architectures, allowing for more nuanced interaction between the text and the knowledge graph. By attending to relevant entities and relations, these models can better capture intricate relationships and perform complex reasoning. However, they can be computationally expensive,

especially with large KGs, and may be sensitive to noise in the graph. Embedding-based approaches are more suitable when computational efficiency and scalability are paramount, while attention-based methods are preferred when high accuracy and complex reasoning are required, even at the cost of increased computational complexity and potential overfitting on noisy KG data. The choice depends on the trade-off between computational cost, model complexity, and the specific requirements of the NLI task.

5.2 Challenges and Limitations

KG-NLI faces several challenges. Knowledge graph incompleteness limits inference capabilities, as missing relationships hinder accurate reasoning. Noise within KGs, stemming from incorrect or outdated information, can lead to flawed conclusions. Scalability poses another hurdle; processing large KGs alongside textual data demands significant computational resources and efficient algorithms. Furthermore, current KG-NLI models often struggle with complex reasoning patterns, particularly those requiring multi-hop inference or nuanced understanding of logical relationships. The ability to effectively integrate symbolic knowledge from KGs with the statistical power of neural networks remains a key area for improvement. Addressing these limitations is crucial for advancing the robustness and applicability of KG-NLI.

6 Future Perspectives

Integrating commonsense knowledge is a promising frontier for NLI, addressing the gap between factual reasoning and everyday understanding. Leveraging structured graphs like ConceptNet and ATOMIC through graph embeddings and attention mechanisms allows models to reason beyond explicit text. Future research should prioritize dynamic fusion of factual and commonsense data using knowledge-aware attention or joint training. Furthermore, exploring reinforcement learning or meta-learning for instance-specific knowledge selection is essential to optimize integration while mitigating potential noise and bias. Explainability and interpretability are crucial for advancing KG-NLI. Current models often operate as black boxes, hindering trust and adoption. Understanding *why* a model arrives at a specific inference is as important as the inference itself. Future research should prioritize techniques for visualizing the knowledge graph paths used in reasoning. Methods like attention mechanisms can highlight relevant nodes and relations, offering insights into the model's decision-making process. Furthermore, generating natural language explanations for the inference chain, detailing how the knowledge graph supports the conclusion, will be vital. This involves identifying the key subgraph G' within the larger knowledge graph G that contributes most to the inference, and articulating the relationship between G' and the premise and hypothesis.

7 Conclusion

Knowledge Graph (KG) -empowered Natural Language Inference (NLI) leverages graph embeddings and attention mechanisms to integrate relational knowledge, with dynamic fusion methods showing superior performance. Despite this potential, significant challenges remain in handling noisy or incomplete KGs and ensuring scalability to large datasets.

To advance the field, future research must prioritize robust knowledge selection and integration strategies. Key focus areas include developing richer KG representations—incorporating relation types and attributes—and enhancing multi-hop reasoning over complex structures. Improving model robustness and generalizability is essential for diverse NLI tasks. Ultimately, these advancements will lead to more accurate and reliable NLU systems, significantly benefiting downstream applications such as question answering and dialogue systems by enabling more intelligent, human-like reasoning.

References

[1] P. Shao, Z. Huang, Y. Qie, X. Xu, L. Qiao, and N. Anwer, "A Review: Knowledge Inference Integrating Large Language Model with Knowledge Graph." Available at SSRN 5097517.

[2] Q. Chen, X. Zhu, Z. H. Ling, and D. Inkpen, "Natural language inference with external knowledge," 2018.

[3] S. Pan, L. Luo, Y. Wang, C. Chen, J. Wang, and X. Wu, "Unifying large language models and knowledge graphs: A roadmap," *IEEE Transactions on Knowledge and Data Engineering*, vol. 36, no. 7, pp. 3580-3599, 2024.

[4] B. Jia, H. Xu, and M. Guo, "Natural Language Inference Using Evidence from Knowledge Graphs," in *International Conference of Pioneering Computer Scientists, Engineers and Educators*, Singapore, 2021, pp. 3-15.

[5] D. V. Vo and P. Do, "Automatically Generating a Dataset for Natural Language Inference Systems from a Knowledge Graph," in *International Conference on Intelligent Systems and Data Science*, Singapore, 2024, pp. 264-276.

[6] A. P. Muharram and A. Purwarianti, "Enhancing natural language inference performance with knowledge graph for COVID-19 automated fact-checking in Indonesian language," arXiv preprint arXiv:2409.00061, 2024.

[7] M. Song, W. Zhao, and E. Haihong, "KGAnet: a knowledge graph attention network for enhancing natural language inference," *Neural Computing and Applications*, vol. 32, no. 18, pp. 14963-14973, 2020.

[8] A. Gajbhiye, N. A. Moubayed, and S. Bradley, "ExBERT: An external knowledge enhanced BERT for natural language inference," in *International Conference on Artificial Neural Networks*, Cham, 2021, pp. 460-472.

[9] P. Schneider, T. Schopf, J. Vladika, M. Galkin, E. Simperl, and F. Matthes, "A decade of knowledge graphs in natural language processing: A survey," in *Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 2022, pp. 601-614.

[10] Z. Yang, Y. Xu, J. Hu, and S. Dong, "Generating knowledge aware explanation for natural language inference," *Information Processing & Management*, vol. 60, no. 2, 103245, 2023.