RULETOPO: NEURO-SYMBOLIC SYNERGY FOR INDUCTIVE KNOWLEDGE GRAPH COMPLETION

Anonymous authors
Paper under double-blind review

ABSTRACT

Knowledge Graph Completion (KGC) infers missing links to enrich incomplete knowledge graphs, categorized into transductive and inductive scenarios. Inductive KGC poses significant challenges, particularly in fully inductive settings with entirely new relations, and still lacks effective modeling paradigms. Recently, rule-based approaches rely on known relation semantics and struggle to generalize to new relations, while relation-graph methods mainly model statistical co-occurrence structures and overlook transferable logical constraints between relations, thus performing poorly when entities and relations are completely unseen. To tackle these issues, we propose **RuleTopo**, a novel Graph Neural Network-based approach that fuses rule learning with a global topology relation graph, supporting fully, semi-, and traditional inductive KGC. Our approach extracts Horn rules and maps them into continuous embeddings, explicitly integrating symbolic rules into the attention mechanism of a relational graph neural network to enable rule-guided topological message passing. Furthermore, we introduce a rule-consistency loss to constrain the logical structure of relation embeddings during training continuously. Experiments show that on the -100/-50/-0 inductive splits of the FB15K-237 and NELL995 datasets, RuleTopo demonstrates higher efficiency and interpretability. This approach demonstrates a scalable, highly generalizable neuro-symbolic cooperative paradigm.

1 INTRODUCTION

Knowledge underpins human intelligence. To emulate human behavior and decision-making, Knowledge Graphs (KGs) have gained attention as tools for structuring, storing, and managing semantic knowledge (Hogan et al., 2021), typically encoding facts as (head entity, relation, tail entity) triplets for retrieval and reasoning. Due to the incompleteness of KGs, Knowledge Graph Completion (KGC) aims to predict missing elements (Arnaout et al., 2021) and is essential for downstream tasks such as question answering (Sui et al., 2025; Jafarzadeh et al., 2025), recommendation systems (Elahi et al., 2025), and intelligent virtual assistants (Men et al., 2025).

KGC methods are commonly categorized into transductive and inductive settings, depending on whether entities and relations appearing at test time are observed during training. While transductive methods, such as TransE (Bordes et al., 2013), has been extensively studied, inductive KGC better reflects real-world scenarios, where new entities and relations continually emerge and must be reasoned about without retraining from scratch. Among inductive settings, fully inductive KGC, where both entities and relations are unseen during training (Teru et al., 2020), poses a particularly fundamental challenge. Unlike new entities, unseen relations introduce novel semantics and interaction patterns that cannot be directly inferred from observed relation embeddings. Despite its practical importance, this setting remains significantly underexplored, and existing solutions often fail to generalize effectively when relation semantics are absent during training.

Existing approaches to inductive KGC can be broadly categorized into symbolic rule-based methods and neural graph-based approaches. However, as shown in Figure 1, both paradigms exhibit inherent limitations when confronted with fully inductive settings involving unseen relations. Rule-based methods leverage logical patterns to enhance reasoning, relying on observed relation semantics and struggling to generalize beyond transductive or partially inductive scenarios. In particular, when relations are entirely unseen during training, the extracted rules cannot be reliably grounded

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

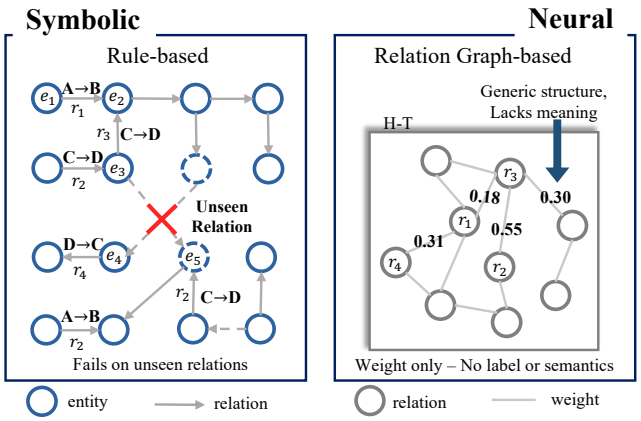

Figure 1: Unifying symbolic rules and neural topology at the relation level.

in learned representations. In contrast, graph-based approaches, such as relation graph neural networks (GNNs), model inductive biases primarily through statistical co-occurrence InGram (Lee et al., 2023) and topological structures (Ding et al., 2025). While these methods generalize better to unseen entities, they overlook the logical constraints that govern relational semantics. As a consequence, neither paradigm alone provides a sufficient inductive bias for reasoning over completely unseen entities and relations.

Our key insight is that symbolic rules and neural topology are not competing sources of inductive bias, but rather complementary signals that should be unified at the relation level. Symbolic rules capture transferable logical constraints among relations that remain valid even when specific entities or relations are unseen, while relation-level topology provides a compact and generalizable abstraction of relational interactions. Crucially, the relation level serves as a natural interface where symbolic semantics can be seamlessly embedded into neural message passing, enabling logical regularities to directly shape representation learning. This perspective motivates a neuro-symbolic synergy that integrates symbolic rules into neural relation graph learning, yielding a principled and effective solution for fully inductive KGC.

Based on this insight, we propose RuleTopo, a GNN-based approach for inductive KGC that tightly couples rule learning with global relation topology. RuleTopo extracts Horn rules and embeds them into continuous representations, which are then integrated into relation GNNs to guide topology-aware message passing, while a rule consistency objective further enforces logical coherence in relation representations. Extensive experiments on FB15K-237 (Toutanova & Chen, 2015) and NELL995 (Xiong et al., 2017) benchmarks under various inductive settings demonstrate that RuleTopo consistently outperforms baselines, achieving substantial gains in fully inductive scenarios.

The main contributions of this work are threefold:

- We introduce RuleTopo, a relation-level neuro-symbolic approach that integrates symbolic rules with neural relation graph learning, to tackle the core difficulty of capturing and generalizing relation semantics when both entities and relations are unseen during training.
- We propose a rule-guided attention mechanism with rule consistency loss, enabling logical constraints to directly guide topology-aware message passing in inductive settings, which together realize neuro-symbolic synergy for inductive representation learning.
- We conduct extensive experiments under multiple inductive settings, demonstrating that RuleTopo consistently outperforms strong baselines.

2 Preliminaries

We first give the problem formulations, then define inductive KGC, relation graphs, and topology patterns.

Problem Formulation. A KG is modeled as a directed multi-relational graph $\mathcal{G} = (\mathcal{E}, \mathcal{R}, \mathcal{T})$, where $\mathcal{E}$ denotes the entity set, $\mathcal{R}$ denotes the relation set, and $\mathcal{T} : \mathcal{E} \times \mathcal{R} \times \mathcal{E} \to \{0, 1\}$ is a binary incidence function indicating whether a fact holds. A triple (h, r, t) is valid if $\mathcal{T}(h, r, t) = 1$, and

KGC (Arnaout et al., 2021) aims to infer such unobserved but valid facts from the observed graph structure.

Inductive KGC. The training and test graphs are disjoint with respect to entities, relations, or both, leading to different levels of inductive generalization. Let $\mathcal{G}_{\mathrm{tr}} = (\mathcal{E}_{\mathrm{tr}}, \mathcal{R}_{\mathrm{tr}}, \mathcal{T}_{\mathrm{tr}})$ and $\mathcal{G}_{\mathrm{te}} = (\mathcal{E}_{\mathrm{te}}, \mathcal{R}_{\mathrm{te}}, \mathcal{T}_{\mathrm{te}})$ denote the training and test graphs. 1) *Fully inductive KGC* (Lee et al., 2023) represents the most challenging setting, where neither entities nor relations are shared between training and test graphs, i.e., $\mathcal{E}_{\mathrm{tr}} \cap \mathcal{E}_{\mathrm{te}} = \emptyset$ and $\mathcal{R}_{\mathrm{tr}} \cap \mathcal{R}_{\mathrm{te}} = \emptyset$. 2) *Semi-inductive KGC* (Lee et al., 2023) further relaxes this assumption by allowing only partial overlap in relations, i.e., $\mathcal{E}_{\mathrm{tr}} \cap \mathcal{E}_{\mathrm{te}} = \emptyset$ and $\mathcal{R}_{\mathrm{tr}} \cap \mathcal{R}_{\mathrm{te}} \neq \emptyset$. 3) *Traditional inductive KGC* (Ding et al., 2025) assumes that all entities are unseen at test time while relations are shared, i.e., $\mathcal{E}_{\mathrm{tr}} \cap \mathcal{E}_{\mathrm{te}} = \emptyset$ and $\mathcal{R}_{\mathrm{tr}} = \mathcal{R}_{\mathrm{te}}$.

Relation Graphs. To abstract relational interactions beyond individual entities, a KG can be transformed into a relation graph in which each node corresponds to a relation in $\mathcal{R}$. Edges in the relation graph represent structural interactions between relations induced by co-occurrence or compositional patterns observed in $\mathcal{T}$ (Lee et al., 2023). Formally, a relation graph is defined as $\mathcal{G}_{\mathcal{R}} = (\mathcal{R}, \mathcal{A})$, where $\mathcal{A}$ denotes adjacency relations between pairs of relations. By modeling topology at the relation level, this abstraction yields a compact and transferable representation of relational structure, which is particularly advantageous for inductive settings involving unseen entities and relations.

Topology Patterns. Following (Chen et al., 2021), we categorize the topological interaction between a pair of relations into six structural patterns: *head-to-tail* (H-T), *tail-to-head* (T-H), *tail-to-tail* (T-T), *head-to-head* (H-H), *parallel* (PARA), and *loop* (LOOP). Given a training triple $(h, r, t) \in \mathcal{T}_{\mathrm{tr}}$, we count co-occurrences of relation pairs (r_1, r_2) under the following configurations: 1) T-H: $h \xrightarrow{r_1} t \xrightarrow{r_2} c$; 2) H-T: $a \xrightarrow{r_2} h \xrightarrow{r_1} t$; 3) T-T: $h \xrightarrow{r_1} t \xleftarrow{r_2} c$ with $c \neq h$; 4) H-H: $a \xleftarrow{r_2} h \xrightarrow{r_1} t$ with $a \neq t$; 5) PARA: $h \xrightarrow{r_1} t$ and $h \xrightarrow{r_2} t$ with $r_1 \neq r_2$; 6) LOOP: $h \xrightarrow{r_1} t$ and $t \xrightarrow{r_2} h$. Each observed instance of these patterns contributes to the corresponding bin-specific edge count between r_1 and r_2.

3 METHODOLOGY

We first give a brief overview and then introduce the proposed new method in detail. Figure 2 illustrates the overview of the proposed approach.

3.1 OVERVIEW

Our approach follows a neuro-symbolic framework that integrates symbolic relational rules with neural relation graph modeling for inductive KGC. It consists of three synergistic components: 1) Given a training graph $\mathcal{G}_{\mathrm{tr}}$, we first extract Horn rules that capture logical regularities among relations and encode them as fixed rule embeddings. 2) In parallel, the KG is abstracted into relation graphs, where each node represents a relation and edges encode relation-level interactions induced by observed triples. The neural component operates on this relation graph to learn topology-aware relation representations via attention-based message passing. Crucially, the extracted rule embeddings are injected into the relation-level attention mechanism, allowing symbolic rules to modulate how information is propagated between related relations. This design enables logical regularities to act as soft structural priors during neural message passing. 3) The resulting relation representations are then used to guide entity-level representation learning and triple scoring in downstream prediction. To further encourage logical coherence, a rule consistency loss is introduced during training to regularize the learned representations with respect to the extracted rules. All components operate primarily at the relation level, which allows the framework to generalize effectively to graphs with unseen entities and relations in fully inductive settings.

3.2 SYMBOLIC – RULE EXTRACTION AND EMBEDDING

Relational regularities in KGs are commonly captured by Horn rules (Horn, 1951), which take the form $r_1(x, y) \wedge r_2(y, z) \wedge \cdots \wedge r_k(\cdot, \cdot) \Rightarrow r_h(x, z)$, where $r_1, \ldots, r_k, r_h \in \mathcal{R}$ denote relations and x, y, z are logical variables that can be instantiated by arbitrary entities in KGs. The conjunction on the left-hand side constitutes the rule body, while the right-hand side represents the rule head. Such

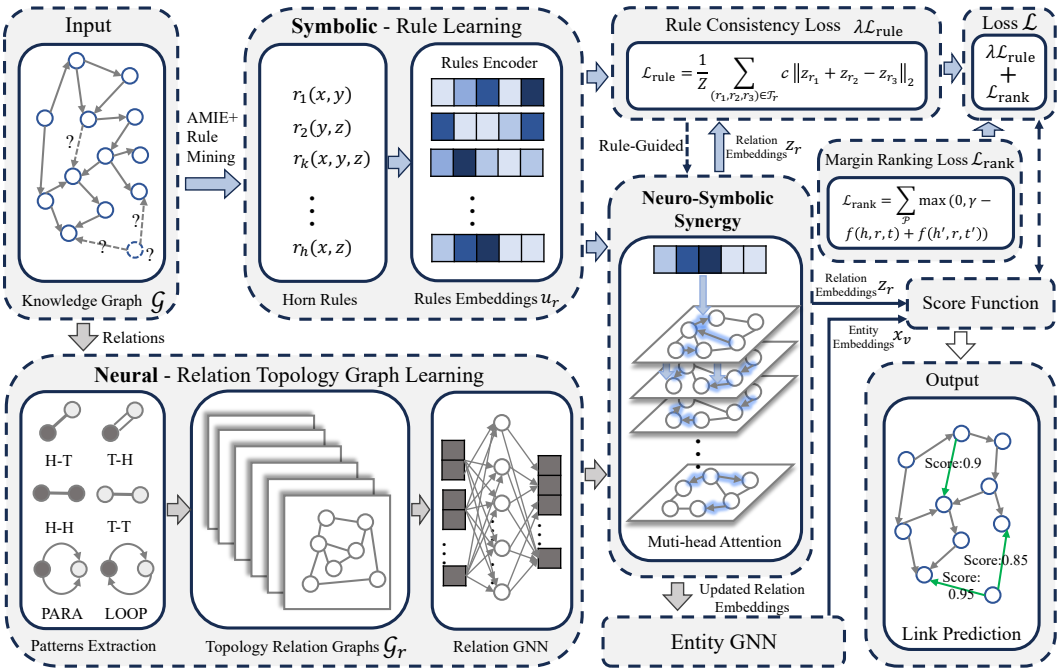

Figure 2: An overview of RuleTopo: A neuro–symbolic framework that integrates Horn rules with neural relation-graph modeling for inductive KGC. Horn rules from $\mathcal{G}$ are encoded as fixed embeddings and injected into relation-level attention; the relation representations guide entity-level learning and triple scoring, with a rule consistency loss enforcing alignment with the rules.

rules encode relation-level dependencies that are independent of specific entity instantiations in the triple set $\mathcal{T}$.

In RuleTopo, symbolic rules are used to capture relation-level regularities that complement topology-driven neural signals. We employ AMIE+ (Galárraga et al., 2015) to mine Horn rules from the observed subgraph of each data split (training/validation). To avoid leaking query facts, rule mining is performed only on the observation set (i.e., message or support triples), excluding evaluation queries. Following standard practice, rule bodies are restricted to at most three atoms ($k \leq 3$).

For each relation r, let $\mathcal{S}_r$ denote the set of mined rules involving r (either in the body or the head). We encode these symbolic rules into a fixed relation-specific rule embedding $\mathbf{u}_r \in \mathbb{R}^{d_r}$ using a set encoder $\mathbf{u}_r = f_{\text{rule}}(\mathcal{S}_r)$, where $f_{\text{rule}}(\cdot)$ aggregates the rule features associated with r. If no rule is mined for relation r (i.e., $\mathcal{S}_r = \varnothing$), we set $\mathbf{u}_r = \mathbf{0}$. These rule embeddings are fixed after mining and serve as symbolic descriptors of relations, which will be integrated with neural relation representations in Sec. 3.4.

3.3 Neural – Relation Graph Neural Network

Relation-level interactions are modeled on a relation graph $\mathcal{G}_{\mathcal{R}} = (\mathcal{R}_{\text{tr}}, \mathcal{A})$, where each node corresponds to a training relation and a directed edge $(r_i \to r_j) \in \mathcal{A}$ indicates that relation r_j is observed to co-occur with r_i under specific global structural patterns in the training graph $\mathcal{G}_{\text{tr}}$ (Lee et al., 2023). Such a graph captures global correlation structures among relations beyond the local neighborhood of entities. To account for heterogeneous interaction patterns, we partition the edges of the relation graphs into topology bins $B = 6$ (Ding et al., 2025), yielding $\{\mathcal{A}^{(b)}\}_{b=1}^{B}$. For undefined relations and entities, we use Glorot initialization (Glorot & Bengio, 2010) to generate initial feature vectors. We then apply GNN-style message passing (Wang et al., 2022) to iteratively update relation and entity representations.

Let $\mathbf{H}^{(0)} \in \mathbb{R}^{|\mathcal{R}_{\text{tr}}| \times d}$ denote the initial relation representations. A relation layer performs attention-based message passing independently within each topology bin. For an edge $(i \to j) \in \mathcal{A}^{(b)}$, we compute an unnormalized attention score $e_{ij}^{(b)}$ (defined in Sec. 3.4), which is normalized by a neighborhood-wise softmax:

$$\beta_{ij}^{(b)} = \frac{\exp\left(e_{ij}^{(b)}\right)}{\sum_{k:(i\to k)\in\mathcal{A}^{(b)}} \exp\left(e_{ik}^{(b)}\right)}. \tag{1}$$

Messages are aggregated to produce a bin-specific update:

$$\mathbf{h}_i^{(b)} = \sum_{j:(i\to j)\in\mathcal{A}^{(b)}} \beta_{ij}^{(b)}\, \mathbf{W}_{\text{msg}} \mathbf{h}_j, \tag{2}$$

where $\mathbf{W}_{\text{msg}}$ is a learnable projection shared across bins. The bin-wise outputs are combined using learnable bin weights:

$$\mathbf{h}_i' = \sum_{b=1}^{B} \pi_b\, \mathbf{h}_i^{(b)}, \tag{3}$$

where $\boldsymbol{\pi} = \mathrm{softmax}(\mathbf{w})$ and $\mathbf{w} \in \mathbb{R}^B$ is trainable. Stacking multiple relation layers with residual connections yields topology-aware relation embeddings.

Given the updated relation embeddings $\mathbf{H} \in \mathbb{R}^{|\mathcal{R}_{\text{tr}}| \times d}$, we update entity representations by relation-conditioned message passing. Let $\mathbf{x}_v \in \mathbb{R}^d$ denote the embedding of entity v. For each outgoing edge (v, r, u), the message from neighbor u to v is modulated by the embedding of relation r (looked up from $\mathbf{H}$) and aggregated over the neighborhood. Following InGram (Lee et al., 2023), we additionally incorporate an explicit self-loop relation (Hamilton et al., 2017) for each entity, whose embedding is computed from incident relations. Stacking multiple entity layers yields context-aware entity embeddings, which are subsequently combined with relation embeddings for triple scoring.

3.4 Neuro – Symbolic Synergy

RuleTopo is motivated by the observation that relation correlations are strongly tied to global topological structures, while symbolic rules provide complementary relation-level regularities. Neuro-symbolic synergy is realized by injecting rule-derived signals into topology-specific relation attention, allowing symbolic and neural cues to jointly determine relation-to-relation message passing.

Let $\mathbf{u}_r \in \mathbb{R}^{d_r}$ denote the fixed rule embedding of relation r (Sec. 3.2). For each relation-graph edge $(i \to j)$, we construct a symmetric rule-induced signal by projecting and combining the rule embeddings of the incident relations:

$$\mathbf{s}_{ij} = \mathbf{W}_{\text{rule}} \mathbf{u}_{r_i} + \mathbf{W}_{\text{rule}} \mathbf{u}_{r_j}, \tag{4}$$

where $\mathbf{W}_{\text{rule}} \in \mathbb{R}^{d \times d_r}$ is a learnable projection. The same projection is applied to both endpoint relations, making the injected signal invariant to swapping the two relations.

Given current neural relation embeddings $\mathbf{h}_i, \mathbf{h}_j \in \mathbb{R}^d$ (Sec. 3.3), topology-specific attention scores are computed by conditioning on both neural and symbolic signals:

$$e_{ij}^{(b)} = \mathbf{a}^\top \sigma(\mathbf{W}_{\text{att}} \left[\mathbf{h}_i \,\|\, \mathbf{h}_j \,\|\, \mathbf{s}_{ij}\right]), \tag{5}$$

where $(i \to j) \in \mathcal{A}^{(b)}$, $b \in \{1, \ldots, B\}$, $\sigma(\cdot)$ is LeakyReLU (Maas et al., 2013), and $\mathbf{W}_{\text{att}}, \mathbf{a}$ are learnable parameters following GATv2 (Brody et al., 2022). Normalization and aggregation follow Eq. (1)–(3).

This design enables symbolic rules to bias relation attention under different topological contexts, while preserving the flexibility of topology-driven neural message passing. The synergy operates purely at the relation level: topology bins are derived from relation co-occurrence structures, and rule embeddings depend only on relations rather than entity identities. In practice, rule injection is lightweight and edge-local. We precompute the projected rule vectors $\tilde{\mathbf{u}}_r = \mathbf{W}_{\text{rule}} \mathbf{u}_r$ for all relations, and form $\mathbf{s}_{ij}$ by a simple symmetric sum. Since $\mathbf{u}_r$ is fixed, back-propagation only updates $\mathbf{W}_{\text{rule}}$, $\mathbf{W}_{\text{att}}$, and $\mathbf{a}$, keeping training stable and computational overhead negligible compared to the base GATv2 computation.

3.5 Optimization Objective

While rule-guided attention injects symbolic priors into relation-level message passing, the learned relation embeddings may still deviate from compositional regularities suggested by mined Horn rules. We therefore introduce a rule consistency loss $\mathcal{L}_{\text{rule}}$ as an auxiliary regularizer.

Let $\mathbf{z}_r \in \mathbb{R}^d$ denote the final relation embedding produced by the relation graph network. From the mined rule set, we construct a collection of relation-composition constraints represented as relation triplets (r_1, r_2, r_3) on the relation graph, indicating that relation r_3 is encouraged to behave like the composition of r_1 and r_2 in the embedding space as $\mathbf{z}_{r_1} + \mathbf{z}_{r_2} \approx \mathbf{z}_{r_3}$.

Let $\mathcal{T}_r$ denote the set of such extracted triplets. Each triplet is weighted by the PCA confidence (Galárraga et al., 2013) of the underlying rule, denoted by $c \geq 0$ (stored in the implementation as a scalar confidence associated with the triplet head). We define $\mathcal{L}_{\text{rule}}$ as:

$$\mathcal{L}_{\text{rule}} = \frac{1}{Z} \sum_{(r_1,r_2,r_3)\in\mathcal{T}_r} c \left\| \mathbf{z}_{r_1} + \mathbf{z}_{r_2} - \mathbf{z}_{r_3} \right\|_2 , \tag{6}$$

where $Z = \text{stopgrad}(\mathbb{E}_{\text{batch}}[c\|\mathbf{z}_{r_1} + \mathbf{z}_{r_2} - \mathbf{z}_{r_3}\|_2]) + \epsilon$ is a self-normalization term, where $\mathbb{E}_{\text{batch}}$ is computed over the current minibatch for stabilizing training.

Besides, we also adopt a margin-based ranking loss with negative sampling follows (Lee et al., 2023):

$$\mathcal{L}_{\text{rank}} = \sum_{((h,r,t),(h',r,t'))\in\mathcal{P}} \max\big(0, \gamma - f(h,r,t) + f(h',r,t')\big), \tag{7}$$

where γ is the margin, $f(\cdot)$ is the model score introduced by DistMult (Yan et al., 2022), and $\mathcal{P}$ consists of pairs between each positive training triple and its sampled negatives.

The overall objective is $\mathcal{L} = \mathcal{L}_{\text{rank}} + \lambda\mathcal{L}_{\text{rule}}$, where λ controls the regularization strength.

4 Experiments

We first give the experimental setup, then compare our approach with baselines on inductive and traditional inductive link prediction. For inductive link prediction, we evaluate performance under both semi-inductive and fully inductive scenarios with respect to relations. Finally, we present comprehensive ablation studies to analyze the contributions of individual components in our approach.

4.1 Experimental Setup

Datasets. We evaluate RuleTopo on two normal public KG benchmarks, NELL995 (Xiong et al., 2017) and FB15K-237 (Toutanova & Chen, 2015). Following the inductive data construction protocol of InGram (Lee et al., 2023), each benchmark is processed into multiple datasets that correspond to different inductive learning scenarios. For more details, see Appendix A.1.

Baselines. We compare the performance of with 16 models: GraIL (Teru et al., 2020), CoMPILE (Mai et al., 2021), RMPI (Geng et al., 2023), CompGCN (Vashishth et al., 2020), NodePiece (Galkin et al., 2022), NeuralLP (Yang et al., 2017), DRUM (Sadeghian et al., 2019), BLP (Daza et al., 2021), QBLP (Ali et al., 2022), NBFNet (Zhu et al., 2021), RED-GNN (Zhang & Yao, 2022), RAILD (Gesese et al., 2022), SE-GNN (Li et al., 2022), GreenKGC (Wang et al., 2023), InGram (Lee et al., 2023) and TARGI (Ding et al., 2025). For more implementation details, see Appendix A.2.

4.2 Inductive Link Prediction

We evaluate RuleTopo on the inductive link prediction task, following the standard protocol of InGram (Lee et al., 2023) and TARGI (Ding et al., 2025), where entities in validation and test graphs are entirely unseen during training. We report Mean Rank (MR, lower is better), Mean Reciprocal Rank (MRR), and Hits@K on four widely used benchmarks: FB-100, NL-100, FB-50, and NL-50. As shown in Table 1, RuleTopo consistently achieves the best or second-best performance across all datasets and metrics. In particular, RuleTopo establishes new state-of-the-art results on FB-100, NL-100, and FB-50 in terms of both MRR and Hits@1, while maintaining strong performance on

Table 1: Inductive link prediction performance on 4 different datasets, where all entities are new, whereas the last digits of each dataset (-100 and -50) indicate the ratio of new relations. The best results are **boldfaced**, and the second-best results are underlined. Our model RuleTopo significantly outperforms the baseline methods in most cases.

	FB-100			NL-100			FB-50			NL-50		
	MRR	Hit@10	Hit@1	MRR	Hit@10	Hit@1	MRR	Hit@10	Hit@1	MRR	Hit@10	Hit@1
CompGCN	0.015	0.020	0.012	0.008	0.015	0.003	0.004	0.007	0.002	0.004	0.008	0.001
NodePiece	0.008	0.006	0.002	0.016	0.018	0.005	0.020	0.046	0.007	0.039	0.086	0.015
SE-GNN	0.005	0.015	0.006	0.022	0.038	0.004	0.011	0.010	0.002	0.005	0.004	0.001
RAILD	0.039	0.060	0.020	0.023	0.047	0.010	-	-	-	-	-	-
NeuralLP	0.026	0.063	0.010	0.079	0.191	0.031	0.082	0.190	0.049	0.099	0.193	0.067
DRUM	0.034	0.072	0.012	0.074	0.135	0.042	0.068	0.133	0.037	0.161	0.330	0.072
BLP	0.022	0.039	0.008	0.022	0.047	0.009	0.070	0.157	0.035	0.037	0.096	0.013
QBLP	0.016	0.025	0.005	0.004	0.004	0.001	0.070	0.140	0.035	0.050	0.111	0.031
NBFNet	0.081	0.166	0.030	0.091	0.203	0.035	0.134	0.265	0.070	0.214	0.334	0.156
GraIL	-	-	-	0.134	0.179	0.125	-	-	-	0.156	0.268	0.099
CoMPILE	-	-	-	0.133	0.218	0.072	-	-	-	0.188	0.309	0.136
RMPI	-	-	-	0.221	0.381	0.141	-	-	-	0.190	0.308	0.102
RED-GNN	0.114	0.260	0.052	0.212	0.377	0.111	0.130	0.252	0.074	0.172	0.290	0.111
GreenKGC	0.002	0.004	0.000	0.006	0.007	0.002	0.140	0.250	0.081	0.197	0.340	0.127
InGram	0.234	0.383	0.166	0.297	0.457	0.215	0.117	0.218	0.067	0.280	0.451	0.193
TARGI	0.251	0.413	0.168	0.315	0.490	0.232	0.153	0.303	0.086	**0.305**	**0.510**	0.202
RuleTopo	**0.272**	**0.433**	**0.194**	**0.334**	**0.497**	**0.248**	**0.166**	**0.319**	**0.090**	0.297	0.460	**0.215**

NL-50. These results demonstrate the effectiveness of integrating symbolic rules with topology-aware relation graph modeling under the inductive setting. Due to scalability issues with enclosing subgraph sampling, we cannot provide the results of GraIL, CoMPILE, and RMPI on FB consistent with the findings in (Lee et al., 2023). The prohibition against new and unknown relations prevents us from offering RAILD results on FB-50 and NL-50.

Compared with InGram, which relies solely on the relation graphs for inductive generalization, RuleTopo yields substantial improvements across all datasets. On FB-100, RuleTopo improves MRR from 0.234 to 0.272 and Hits@1 from 0.166 to 0.194, indicating a more accurate ranking of unseen entity pairs. Similar gains are observed on NL-100, where MRR increases from 0.297 to 0.334 and Hits@1 from 0.215 to 0.248. These improvements suggest that purely relation propagation, while effective, is insufficient to capture relation semantics in inductive scenarios fully. By injecting symbolic rule signals into relation-level message passing, RuleTopo provides additional inductive bias that helps disambiguate relation interactions when encountering unseen entities.

RuleTopo also outperforms TARGI, which represents a strong topology-aware relation graph baseline. For example, on FB-100, RuleTopo improves MRR from 0.251 to 0.272 and Hits@1 from 0.168 to 0.194. On NL-100, RuleTopo achieves an MRR of 0.334 and Hits@1 of 0.248, surpassing TARGI's 0.315 / 0.232 in Hits@1 while remaining competitive in higher-rank metrics. Notably, both TARGI and RuleTopo operate at the relation-graph level and are entity-agnostic, but RuleTopo further enhances relation representations by incorporating symbolic regularities. This comparison highlights that symbolic rules provide complementary information beyond topological co-occurrence patterns, even when strong topology modeling is already in place.

Overall, these results demonstrate that combining symbolic rule embeddings with topology relation graph neural networks leads to more robust and generalizable relation learning. RuleTopo benefits from both sources of inductive bias: topology captures global structural correlation patterns, while symbolic rules encode higher-order relational regularities. Their integration is particularly effective in inductive link prediction, which can be relied upon.

4.3 Traditional Inductive Link Prediction

We further evaluate RuleTopo under the traditional inductive link prediction setting, where the training and test graphs share the same relation set, but entities in the test graph are unseen during training. Following prior work, we report results on NELL995-v1 and NL-0, using MRR and Hits@K as evaluation metrics. As shown in Table 4.3, RuleTopo achieves the best performance across all

metrics on NELL995-v1, with MRR 0.762, Hits@10 0.910, and Hits@1 0.650, substantially outperforming strong inductive baselines such as InGram (Lee et al., 2023) and TARGI (Ding et al., 2025).

We observe that the performance gains of RuleTopo are less pronounced on NL-0. This can be attributed to dataset characteristics: NL-0 exhibits a large discrepancy between the numbers of relations and triples (Table 4), leading to sparse relation co-occurrence patterns and relatively weak rule correlations. As a result, many mined rules on NL-0 are less informative at the relation level, which limits the effectiveness of rule-guided attention in this setting. In contrast, RuleTopo shows substantially larger improvements on NELL-995-v1. This dataset contains a much smaller relation set but a considerably larger number of triples, resulting in denser relational interactions and more closed relational paths. Such structural properties facilitate the mining of stronger and more consistent Horn rules, which in turn provide more effective symbolic guidance for relation-level message passing. These observations suggest that RuleTopo particularly benefits datasets with compact relation vocabularies and rich relational connectivity. Overall, integrating symbolic rule embeddings with topology-aware relation graph modeling not only benefits fully inductive settings but also generalizes well to traditional inductive scenarios.

Table 2: Traditional inductive link prediction.

	NELL995-v1			NL-0		
Model	MRR	Hit@10	Hit@1	MRR	Hit@10	Hit@1
SE-GNN	0.029	0.038	0.007	0.016	0.018	0.009
RAILD	0.065	0.221	0.008	0.053	0.115	0.020
BLP	0.165	0.475	0.054	0.043	0.108	0.015
GreenKGC	0.188	0.351	0.092	0.179	<u>0.418</u>	0.062
CompGCN	0.300	0.740	0.006	0.008	0.009	0.001
QBLP	0.333	0.580	0.244	0.073	0.151	0.016
RMPI	0.479	0.567	0.438	0.221	0.335	0.155
CoMPILE	0.481	0.560	0.402	0.222	0.361	0.145
GraIL	0.489	0.591	0.400	0.196	0.341	0.122
DRUM	0.520	0.738	0.407	0.198	0.333	0.128
RED-GNN	0.552	0.705	0.483	0.221	0.358	0.145
NeuralLP	0.559	0.800	0.413	0.177	0.341	0.099
NBFNet	0.618	0.850	0.511	**0.260**	**0.422**	**0.167**
NodePiece	0.678	<u>0.867</u>	0.566	0.099	0.225	0.041
InGram	0.688	0.727	0.533	0.240	0.411	0.156
TARGI	<u>0.699</u>	0.858	<u>0.580</u>	<u>0.243</u>	0.410	<u>0.160</u>
RuleTopo	**0.762**	**0.910**	**0.650**	0.234	0.414	0.149

4.4 Ablation Study

We conduct an ablation study on FB-100 and NL-100 to examine the contribution of each component in RuleTopo. As shown in Table 4.4, removing any major component consistently degrades performance, confirming that the overall effectiveness of RuleTopo arises from the joint design of topology modeling and symbolic rule integration. Specifically, removing topology modeling results in the largest performance drop, highlighting its role as the backbone of relation representation learning. Disabling rule injection also leads to a clear degradation, especially on NL-100, demonstrating that symbolic rules provide complementary guidance beyond purely topology-driven message passing. In contrast, removing the rule loss causes a smaller but consistent decrease, indicating that it mainly serves as a regularizer to stabilize rule embeddings.

Table 3: Ablation study of RuleTopo.

	FB-100			NL-100		
	MRR	Hit@10	Hit@1	MRR	Hit@10	Hit@1
w/o rule loss	0.244	0.408	0.163	0.328	0.457	0.230
w/o topology	0.233	0.383	0.160	0.300	0.466	0.212
w/o rule injection	0.252	0.420	0.161	0.314	0.489	0.217
w/o mean agg.	0.244	0.384	0.170	0.279	0.467	0.178
RuleTopo	**0.253**	**0.429**	**0.172**	**0.334**	**0.497**	**0.248**

We also observe that replacing mean aggregation with alternative schemes introduces additional performance variance. This suggests that the aggregation strategy affects the stability of relation-level message passing in inductive settings. Overall, the ablation results validate that each component of RuleTopo contributes positively to the final performance, and that symbolic rules act as a lightweight yet effective enhancement to topology-aware relation modeling.

5 Related Works

Knowledge Graph Completion. Representative approaches include translation-based models, i.e., TransE (Bordes et al., 2013) and their extensions, i.e., RotatE (Sun et al., 2019), which learn fixed embeddings for entities and relations and achieve strong performance when entity identities are shared between training and test graphs. However, these methods fail to generalize to inductive scenarios involving unseen entities or relations. To address this limitation, recent work has focused on inductive KGC, i.e., NBFNET (Zhu et al., 2021), where models are required to generalize beyond observed entity sets. Subgraph-based approaches, i.e., GraIL (Teru et al., 2020) and (Hwang et al., 2025), extract local enclosing subgraphs around query entities and apply GNNs to infer missing links, enabling generalization to unseen entities. Other methods adopt relation-centric approaches, i.e., NodePiece (Galkin et al., 2022), decoupling prediction from specific entity identities. While effective in traditional inductive settings with known relations, these methods typically struggle when relations are also unseen, limiting their applicability in fully inductive scenarios.

Symbolic Rule-Based Reasoning. Symbolic reasoning over KGs has long leveraged logical rules to capture relational regularities and dependencies. Horn rules (Horn, 1951), in particular, provide a compact and interpretable representation of relational implications and have been widely studied in the context of rule mining and logical inference. Such rules are inherently entity-agnostic, making them potentially transferable across graphs with unseen entities. Several methods incorporate symbolic rules into KGC by mining rules from observed triples, i.e., AMIE (Galárraga et al., 2013) and AMIE+ (Galárraga et al., 2015), and using them for logical inference or score refinement, i.e., NeuralLP (Yang et al., 2017) and DRUM (Sadeghian et al., 2019). Some works follow up GraIL, i.e., CoMPILE (Mai et al., 2021) and SNRI (Xu et al., 2022), have focused on rule learning, and (Liu et al., 2017) proposes GNN-based encoding capturing logical rules. However, most rule-based or rule-augmented methods are limited to transductive or semi-inductive settings, and often apply rules as post-hoc constraints or separate inference modules, rather than integrating them into neural representation learning in an end-to-end manner.

Neural Topology-Based Methods. Another line of work investigates topology-aware inductive reasoning by abstracting structural patterns in KGs. Relation graph–based methods construct global graphs over relations and apply GNNs to model interactions among relations, capturing higher-order structural dependencies beyond individual entities, i.e., InGram (Lee et al., 2023). Such approaches have demonstrated strong inductive generalization, especially in settings involving unseen entities and relations. Despite their effectiveness, existing topology-based methods typically rely purely on neural message passing and overlook symbolic relational regularities that can provide complementary inductive biases, i.e., TARGI (Ding et al., 2025), Ultra (Galkin et al., 2024), and AdaRPT (Xie et al., 2025). In contrast, our work bridges symbolic rule and neural topology modeling by injecting rule-derived signals directly into topology-aware attention, enabling joint exploitation of logical and structural regularities for fully inductive KGC.

6 Conclusion

In this paper, we propose RuleTopo, a neuro-symbolic approach for inductive KGC that integrates symbolic relational rules with topology-aware relation graph modeling. Unlike prior inductive KGC methods, RuleTopo explicitly captures global relational regularities through Horn rules and injects them into relation-layer attention as soft structural priors. By operating primarily at the relation level, RuleTopo naturally generalizes to fully inductive settings where both entities and relations are unseen during inference. The proposed rule- and topology-guided attention mechanism enables effective information propagation across related relations, while the rule consistency loss further regularizes learned representations toward logical coherence. Extensive experiments on standard inductive benchmarks demonstrate that our work consistently outperforms strong baselines.

AI USE STATEMENT

In this work, we used generative AI tools to assist with language polishing, improving the clarity and presentation of the manuscript, and providing limited programming support during implementation and debugging. We did not use generative AI tools to generate or fabricate experimental results, to analyze data without author verification, or to produce the core scientific contributions, including the proposed methodology, theoretical claims, and experimental conclusions. Other categories requiring disclosure are not applicable to this work. We reviewed and verified all AI-assisted text and code. In particular, all technical descriptions, mathematical formulations, experimental settings, results, and conclusions were checked by the authors; any AI-assisted code was inspected, integrated, and tested by the authors. We take full responsibility for the final content of this work, including all text, claims, and artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

We support reproducibility by providing a detailed description of RuleTopo, including its rule-guided topological message-passing mechanism and rule-consistency training objective, in Sections 3 and 4. The experimental protocol, benchmark construction, baseline comparisons, and evaluation metrics are described in Section 4. We use the publicly available FB15K-237 and NELL995 benchmarks, following the inductive data construction protocol of prior work. Appendix A provides dataset statistics and further implementation details. Upon publication, we will release the complete source code, preprocessing scripts, dataset splits, configuration files, and instructions for reproducing all reported experiments.

REFERENCES

Mehdi Ali, Max Berrendorf, Mikhail Galkin, Veronika Thost, Tengfei Ma, Volker Tresp, and Jens Lehmann. Improving inductive link prediction using hyper-relational facts. In *Proceedings of the 31st International Joint Conference on Artificial Intelligence*, pp. 5259–5263, 2022.

Hiba Arnaout, Simon Razniewski, Gerhard Weikum, and Jeff Z. Pan. Negative statements considered useful. *J. Web Semant.*, 71:100661, 2021.

Antoine Bordes, Nicolas Usunier, Alberto García-Durán, Jason Weston, and Oksana Yakhnenko. Translating embeddings for modeling multi-relational data. In *Advances in Neural Information Processing Systems*, pp. 2787–2795, 2013.

Shaked Brody, Uri Alon, and Eran Yahav. How attentive are graph attention networks? In *Proceedings of the 10th International Conference on Learning Representations*, 2022.

Jiajun Chen, Huarui He, Feng Wu, and Jie Wang. Topology-aware correlations between relations for inductive link prediction in knowledge graphs. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence*, pp. 6271–6278, 2021.

Daniel Daza, Michael Cochez, and Paul Groth. Inductive entity representations from text via link prediction. In *Proceedings of the 2021 ACM Web Conference*, pp. 798–808, 2021.

Ling Ding, Lei Huang, Zhizhi Yu, Di Jin, and Dongxiao He. Towards global-topology relation graph for inductive knowledge graph completion. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 11581–11589, 2025.

Ehsan Elahi, Sajid Anwar, Mousa Al-kfairy, Joel JPC Rodrigues, Alladoumbaye Ngueilbaye, Zahid Halim, and Muhammad Waqas. Graph attention-based neural collaborative filtering for item-specific recommendation system using knowledge graph. *Expert Systems with Applications*, 266: 126133, 2025.

Luis Galárraga, Christina Teflioudi, Katja Hose, and Fabian M. Suchanek. Fast rule mining in ontological knowledge bases with amie+. *The VLDB Journal*, 24(6):707–730, 2015.

Luis Antonio Galárraga, Christina Teflioudi, Katja Hose, and Fabian Suchanek. Amie: association rule mining under incomplete evidence in ontological knowledge bases. In *Proceedings of the 22nd International Conference on World Wide Web*, pp. 413–422, 2013.

Mikhail Galkin, Etienne G. Denis, Jiapeng Wu, and William L. Hamilton. Nodepiece: Compositional and parameter-efficient representations of large knowledge graphs. In *Proceedings of the 10th International Conference on Learning Representations*, 2022.

Mikhail Galkin, Xinyu Yuan, Hesham Mostafa, Jian Tang, and Zhaocheng Zhu. Towards foundation models for knowledge graph reasoning. In *The Twelfth International Conference on Learning Representations*, 2024.

Yuxia Geng, Jiaoyan Chen, Jeff Z. Pan, Mingyang Chen, Song Jiang, Wen Zhang, and Huajun Chen. Relational message passing for fully inductive knowledge graph completion. In *Proceedings of the 39th IEEE International Conference on Data Engineering*, pp. 1221–1233, 2023.

Genet Asefa Gesese, Harald Sack, and Mehwish Alam. RAILD: towards leveraging relation features for inductive link prediction in knowledge graphs. In *Proceedings of the 11th International Joint Conference on Knowledge Graphs*, pp. 82–90, 2022.

Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics*, pp. 249–256, 2010.

William L. Hamilton, Zhitao Ying, and Jure Leskovec. Inductive representation learning on large graphs. In *Advances in Neural Information Processing Systems*, pp. 1024–1034, 2017.

Aidan Hogan, Eva Blomqvist, Michael Cochez, Claudia d'Amato, Gerard De Melo, Claudio Gutierrez, Sabrina Kirrane, José Emilio Labra Gayo, Roberto Navigli, Sebastian Neumaier, et al. Knowledge graphs. *ACM Computing Surveys*, 54(4):1–37, 2021.

Alfred Horn. On sentences which are true of direct unions of algebras1. *The Journal of Symbolic Logic*, 16(1):14–21, 1951.

Minsung Hwang, Jaejun Lee, and Joyce Jiyoung Whang. Stability and generalization capability of subgraph reasoning models for inductive knowledge graph completion. In *Forty-second International Conference on Machine Learning*, 2025.

Parastoo Jafarzadeh, Faezeh Ensan, Mahdiyar Ali Akbar Alavi, and Fattane Zarrinkalam. A knowledge graph embedding model for answering factoid entity questions. *ACM Trans. Inf. Syst.*, 43 (2), 2025. ISSN 1046-8188.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations*, 2015.

Jaejun Lee, Chanyoung Chung, and Joyce Jiyoung Whang. Ingram: Inductive knowledge graph embedding via relation graphs. In *Proceedings of International Conference on Machine Learning*, pp. 18796–18809, 2023.

Ren Li, Yanan Cao, Qiannan Zhu, Guanqun Bi, Fang Fang, Yi Liu, and Qian Li. How does knowledge graph embedding extrapolate to unseen data: A semantic evidence view. In *Proceedings of the 36th AAAI Conference on Artificial Intelligence*, pp. 5781–5791, 2022.

Hanxiao Liu, Yuexin Wu, and Yiming Yang. Analogical inference for multi-relational embeddings. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70, pp. 2168–2178, 2017.

Andrew L Maas, Awni Y Hannun, Andrew Y Ng, et al. Rectifier nonlinearities improve neural network acoustic models. In *Proceedings of International Conference on Machine Learning*, pp. 3, 2013.

Sijie Mai, Shuangjia Zheng, Yuedong Yang, and Haifeng Hu. Communicative message passing for inductive relation reasoning. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence*, pp. 4294–4302, 2021.

ChangHao Men, Yu Han, Ping Wang, JianYuan Tao, and Cheng-Geng Huang. The interpretable reasoning and intelligent decision-making based on event knowledge graph with llms in fault diagnosis scenarios. *IEEE Transactions on Instrumentation and Measurement*, 74:1–16, 2025.

Ali Sadeghian, Mohammadreza Armandpour, Patrick Ding, and Daisy Zhe Wang. DRUM: end-to-end differentiable rule mining on knowledge graphs. In *Advances in Neural Information Processing Systems*, pp. 15321–15331, 2019.

Yuan Sui, Yufei He, Zifeng Ding, and Bryan Hooi. Can knowledge graphs make large language models more trustworthy? an empirical study over open-ended question answering. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics*, pp. 12685–12701, 2025.

Zhiqing Sun, Zhi-Hong Deng, Jian-Yun Nie, and Jian Tang. Rotate: Knowledge graph embedding by relational rotation in complex space. In *Proceedings of the 7th International Conference on Learning Representations*, 2019.

Komal K. Teru, Etienne G. Denis, and William L. Hamilton. Inductive relation prediction by subgraph reasoning. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119, pp. 9448–9457, 2020.

Kristina Toutanova and Danqi Chen. Observed versus latent features for knowledge base and text inference. In *Proceedings of the 3rd Workshop on Continuous Vector Space Models and their Compositionality*, pp. 57–66, 2015.

Shikhar Vashishth, Soumya Sanyal, Vikram Nitin, and Partha P. Talukdar. Composition-based multi-relational graph convolutional networks. In *8th International Conference on Learning Representations*, 2020.

Yuncheng Wang, Xiou Ge, Bin Wang, and C.-C. Jay Kuo. Greenkgc: A lightweight knowledge graph completion method. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, pp. 10596–10613, 2023.

Zhen Wang, Chunjiang Mu, Shuyue Hu, Chen Chu, and Xuelong Li. Modelling the dynamics of regret minimization in large agent populations: a master equation approach. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*, pp. 534–540, 2022.

Zhiwen Xie, Zhuo Zhao, Jinjin Ma, Guangyou Zhou, and Jimmy Xiangji Huang. Adarpt: An adaptive rule pattern transfer model for fully inductive knowledge graph reasoning. In *Proceedings of the 48th International ACM SIGIR Conference*, pp. 1207–1217, 2025.

Wenhan Xiong, Thien Hoang, and William Yang Wang. DeepPath: A reinforcement learning method for knowledge graph reasoning. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pp. 564–573, 2017.

Xiaohan Xu, Peng Zhang, Yongquan He, Chengpeng Chao, and Chaoyang Yan. Subgraph neighboring relations infomax for inductive link prediction on knowledge graphs. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*, pp. 2341–2347, 2022.

Zuoyu Yan, Tengfei Ma, Liangcai Gao, Zhi Tang, and Chao Chen. Cycle representation learning for inductive relation prediction. In *International Conference on Machine Learning*, volume 162, pp. 24895–24910, 2022.

Fan Yang, Zhilin Yang, and William W. Cohen. Differentiable learning of logical rules for knowledge base reasoning. In *Advances in Neural Information Processing Systems*, pp. 2319–2328, 2017.

Yongqi Zhang and Quanming Yao. Knowledge graph reasoning with relational digraph. In *Proceedings of the 22th ACM Web Conference*, pp. 912–924, 2022.

Zhaocheng Zhu, Zuobai Zhang, Louis-Pascal A. C. Xhonneux, and Jian Tang. Neural bellman-ford networks: A general graph neural network framework for link prediction. In *Advances in Neural Information Processing Systems*, pp. 29476–29490, 2021.

A Appendix

A.1 Dataset statistics

The dataset suffix "-50" indicates that 50% of relations in the inference graph are new, while "-100" indicates that all relations are new. For all datasets, we follow the standard inductive assumption of (Teru et al., 2020), where entities in the inference graph are entirely disjoint from those in the training graph. For more details, see Table 4.

Table 4: Dataset statistics, including entities, relations, triplets, mean rule confidence (MCon), and standard deviation (StD). NL-v1 is short for NELL-995-v1.

Dataset	Split	$\lvert\mathcal{E}\rvert$	$\lvert\mathcal{R}\rvert$	$\lvert\mathcal{T}\rvert$	MCon	StD
FB-100	$\mathcal{G}_{\mathrm{tr}}$	4,659	134	62,809	0.66	0.35
	$\mathcal{G}_{\mathrm{va}}$	2,624	77	6,987	0.70	0.39
NL-100	$\mathcal{G}_{\mathrm{tr}}$	1,258	55	7,832	0.70	0.33
	$\mathcal{G}_{\mathrm{va}}$	1,709	53	2,378	0.56	0.39
FB-50	$\mathcal{G}_{\mathrm{tr}}$	5,190	153	91,571	0.63	0.32
	$\mathcal{G}_{\mathrm{va}}$	4,445	205	11,636	0.80	0.34
NL-50	$\mathcal{G}_{\mathrm{tr}}$	4,396	106	17,578	0.59	0.36
	$\mathcal{G}_{\mathrm{va}}$	2,335	119	2,576	0.42	0.12
NL-v1	$\mathcal{G}_{\mathrm{tr}}$	3,103	14	4,687	0.70	0.33
	$\mathcal{G}_{\mathrm{va}}$	3,103	14	4,687	0.70	0.33
NL-0	$\mathcal{G}_{\mathrm{tr}}$	1,814	134	7,796	0.74	0.33
	$\mathcal{G}_{\mathrm{va}}$	2,026	112	2,287	–	–

A.2 Implementation Details

RuleTopo is implemented in PyTorch. All neural components are trained using the Adam optimizer (Kingma & Ba, 2015) with default momentum parameters. The logical rules are mined using an efficient rule mining tool AMIE+ (Galárraga et al., 2015). Relation and rule embeddings are initialized randomly and projected into a shared latent space of dimension d. The relation GNN consists of multiple stacked relation layers with residual connections, and attention-based message passing is performed separately within each topology bin. During training, we optimize the overall objective composed of the link prediction loss and the rule consistency regularization. Negative samples are generated by corrupting either the head or tail entity of each positive triple. Models are trained for a fixed number of epochs, and the checkpoint with the best validation performance is selected for evaluation. All experiments are performed on an NVIDIA RTX 3090 GPU with 24 GB of memory.