

## APPENDIX

### Anonymous authors

Paper under double-blind review

In this appendix, we first give a summary of all notations used in this paper. Then we give an overview of the proposed **HiRe** framework. We also provide the detailed results of ablation study and hyper-parameter sensitivity test on Nell-One.

### A NOTATIONS

The notations used in this paper are summarized in Table 1.

Table 1: Notations used in this paper.

Symbol	Description
$\mathcal{G}$	knowledge graph
$\mathcal{E}, \mathcal{R}, \mathcal{TP}$	entity, relation and triplet sets of a knowledge graph
$h, t$	head entity, tail entity
$r$	relation
$(h, r, t)$	factual triplet
$\mathbf{h}, \mathbf{r}, \mathbf{t}$	embeddings of $h, r$ and $t$
$\mathcal{T}_r$	few-shot task corresponding to relation $r$
$\mathcal{S}_r$	reference set corresponding to relation $r$
$\mathcal{Q}_r$	query set corresponding to relation $r$
$\mathcal{C}_{(\tilde{h}_j, r)}$	candidate set for the potential tail entity of $(\tilde{h}_j, r, ?)$
$\mathcal{N}_e$	neighborhood of entity $e$
$\mathcal{C}_{(h, r, t)}$	context of triplet $(h, r, t)$

### B OVERVIEW OF THE PROPOSED HiRe FRAMEWORK

Figure 1 shows an overview of our proposed hierarchical relational learning (HiRe) framework.

### C ABLATION STUDY ON NELL-ONE

Table 2 illustrates the contribution of each component to the overall performance on Nell-One. As discussed in Section 5.5, each component in our proposed HiRe framework plays an important role in few-shot KG completion. The removal of any component leads to performance drops in terms of all evaluation metrics.

Table 2: Ablation study of HiRe under 3-shot and 5-shot settings on Nell-One.

Ablation on ↓	Components			3-shot				5-shot			
	MTransD	MRL	Context	MRR	Hits@10	Hits@5	Hits@1	MRR	Hits@10	Hits@5	Hits@1
HiRe	✓	✓	✓	0.300	0.499	0.425	0.199	0.306	0.520	0.439	0.207
w/o MTransD	✗	✓	✓	0.295	0.491	0.420	0.193	0.302	0.515	0.428	0.202
w/o MRL-AVG	✓	✗	✓	0.282	0.467	0.382	0.185	0.286	0.466	0.394	0.188
w/o MRL-LSTM	✓	✗	✓	0.280	0.490	0.410	0.168	0.285	0.459	0.390	0.185
w/o Context	✓	✓	✗	0.290	0.482	0.401	0.186	0.295	0.489	0.418	0.197

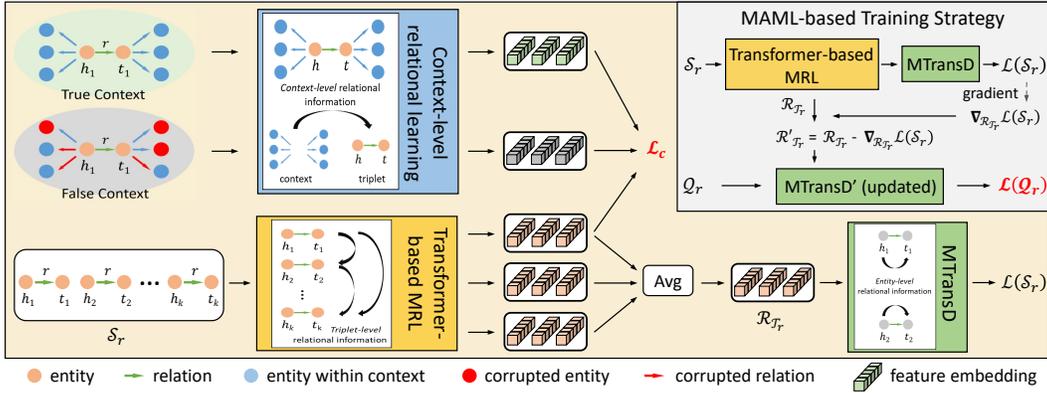


Figure 1: An overview of **HiRe** framework composed of three key components. (1) Contrastive learning based context-level relational learning; (2) Transformer-based triplet-level relational learning; (3) Meta representation based entity-level relational learning. Given a target relation  $r$  and its corresponding reference set  $\mathcal{S}_r$  and query set  $\mathcal{Q}_r$ , we employ a contrastive loss  $\mathcal{L}_c$  between the true/false contexts and the anchor triplet (take  $(h_1, r, t_1)$  as an example) via our proposed contrastive learning based context-level relational learning method. The meta representation of the target relation  $\mathcal{R}_{\mathcal{T}_r}$  is learned by our Transformer-based meta relation learner (MRL), capturing pairwise triplet-level relational information. Lastly, MTransD refines the learned meta relation representation at the entity level constrained by  $\mathcal{L}(\mathcal{S}_r)$ . The whole learning framework is optimized by a MAML-based training strategy.

## D HYPER-PARAMETER SENSITIVITY STUDY ON NELL-ONE

The sensitivity test on hyper-parameter  $N$  and  $\lambda$  is shown in Figure 2 and Figure 3. As illustrated in Figure 2, as the number of false contexts increases, the performance of HiRe drops slightly because its model training converges to a sub-optimal state. In addition, the best  $\lambda$  value is also 0.05 on Nell-One. The overall results are consistent with the conclusion we draw from the results on Wiki-One in Section 5.6.

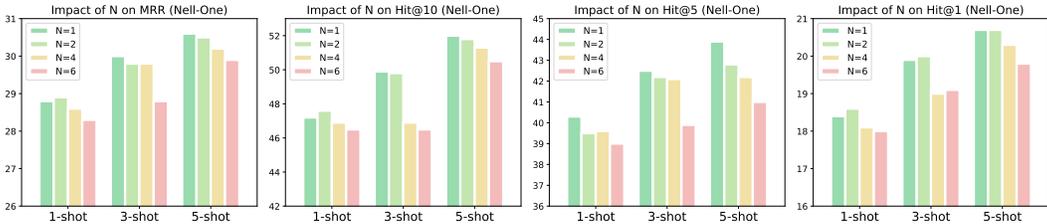


Figure 2: Hyper-parameter sensitivity study of the number of false contexts on Nell-One.

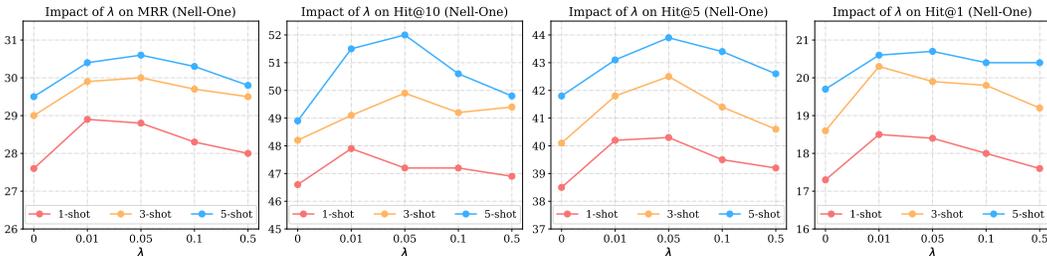


Figure 3: The impact of different  $\lambda$  in Eq. 21 on Nell-One.  $\lambda = 0$  means that we remove the contrastive learning based context-level relational learning.