Neural-Symbolic Entangled Framework for Complex Query Answering

Zezhong Xu[1^], Wen Zhang[1^], Peng Ye[1], Hui Chen[3], Huajun Chen[1,2] *

[1] Zhejiang University & AZFT Joint Lab for Knowledge Engine, China
[2] Hangzhou Innovation Center, Zhejiang University
[3] Alibaba Group

* Corresponding Author
^ Equal Contribution

Introduction

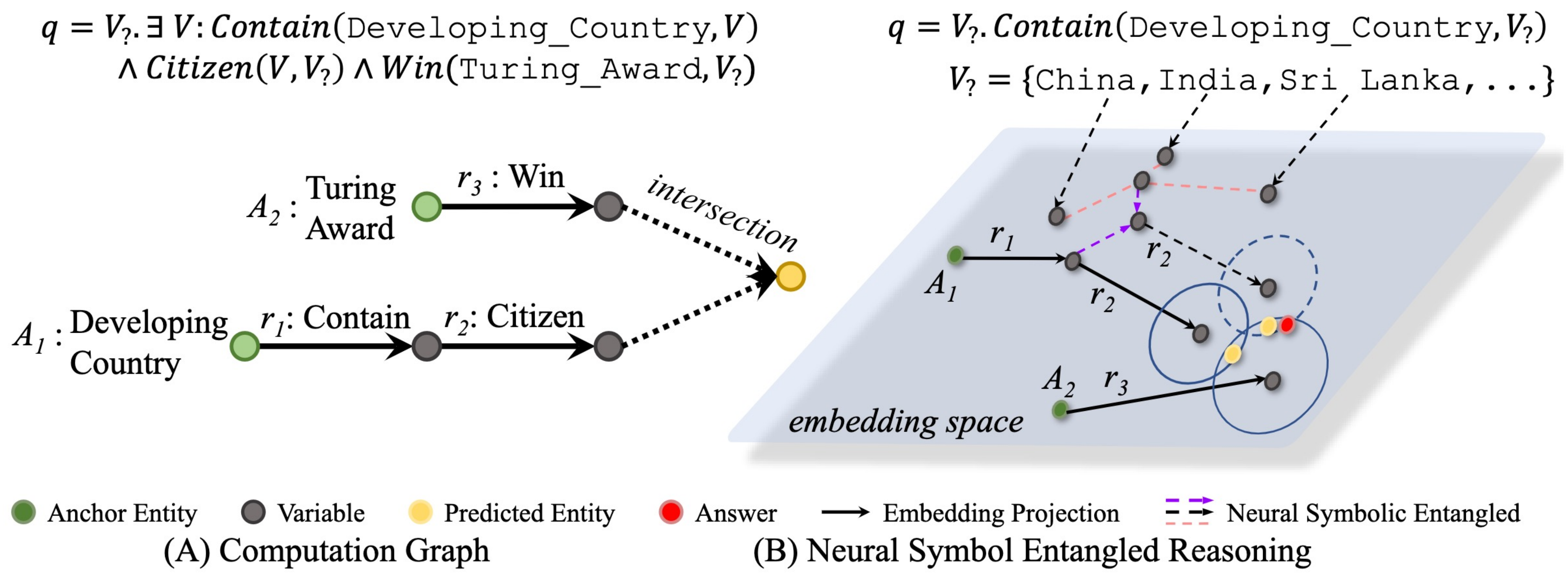


Figure 1: An example of answering a complex graph query by using the ENeSy. (A): FOL query and its computation graph for the question *'Who won the Turing Award in developing countries?'*. (B): ENeSy uses neural and symbolic ways to handle projection separately, and the results are entangled to enhance each other to alleviate the problem of cascading error and incompleteness of KG. The logic operator $\wedge$, $\vee$, and $\neg$ are supported with symbolic reasoning.

Motivation

- Previous works mainly concentrate on the target answers, **ignoring intermediate entities**' usefulness, which is essential for relieving the cascading error problem in logical query answering.

- These methods are usually designed with their own geometric or distributional embeddings to handle logical operators like union, intersection, and negation, with the sacrifice of the accuracy of the basic operator -- projection, and they could not **generalize** other embedding methods to their models.

- We propose a **N**eural and **Sy**mbolic **E**ntangled framework, **ENeSy,** for logical query answering, which enables the neural and symbolic reasoning to **enhance each other** to alleviate the cascading error and KG incompleteness.

Introduction

Overall Aim

- Propose a neural and symbolic entangled framework that could alleviate the problem of cascading error.

Specific Objectives

- Design a model that can enable neural and symbolic reasoning support each other.
- Generalizing existing embedding methods to complex query answering.
- Train the model with link prediction task.

NEURAL INFORMATION PROCESSING SYSTEMS

Logical Operator

- **Relational Projection:** Given an entity set $\mathcal{S} \subseteq \mathcal{V}$ and a relation $r \in \mathcal{R}$, the projection operator return a new entity set $\mathcal{S}'$ that contains the entities related to at least one of entity in $\mathcal{S}$: $\mathcal{S}' = \{e' \in \mathcal{V} | \exists r(e, e'), e \in \mathcal{S}\}$.
- **Intersection:** Given sets of entities $\{S_1, S_2, \ldots, S_n\}$ where $S_i \subseteq \mathcal{V}$, the intersection operator returns the intersection of these sets $\bigcap_{i=1}^{n} S_i$.
- **Union:** Given sets of entities $\{S_1, S_2, \ldots .S_n\}$ where $S_i \subseteq \mathcal{V}$, the union operator returns the union of these sets $\bigcup_{i=1}^{n} S_i$.
- **Complement:** Given a set of entities $\mathcal{S}$, the complement operator returns its complement $\mathcal{S}' = \mathcal{V} - \mathcal{S}$.

Method

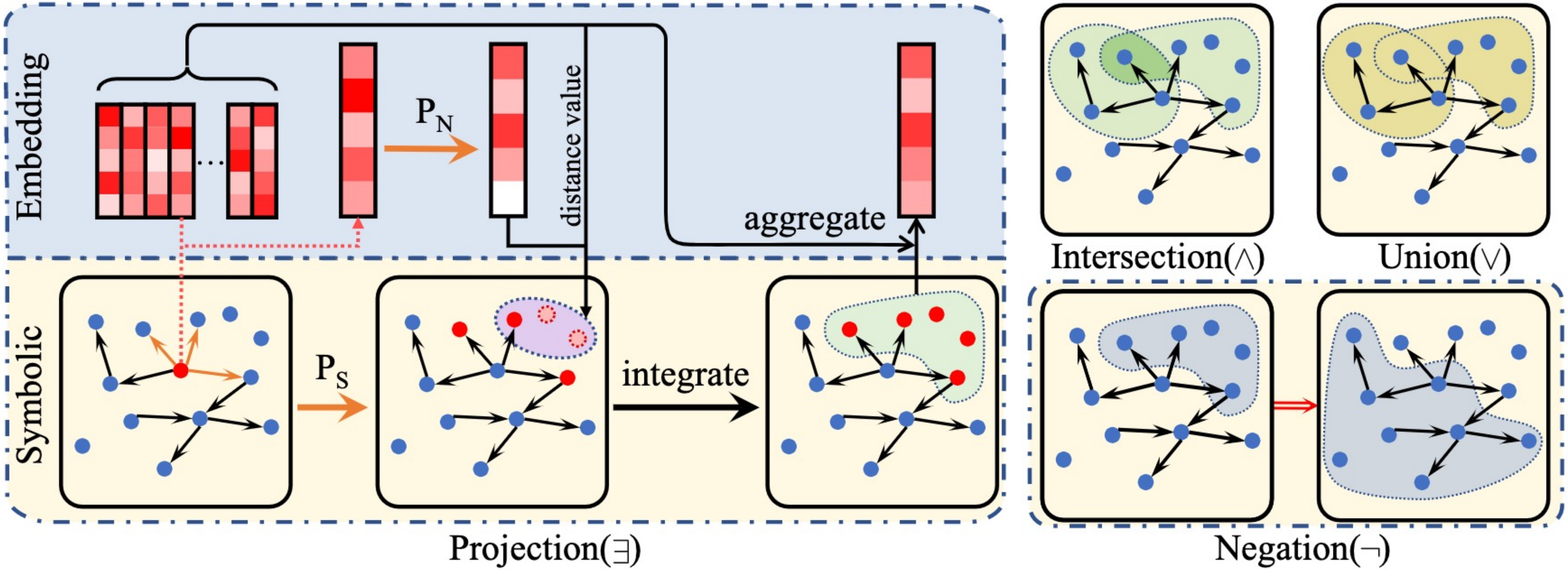
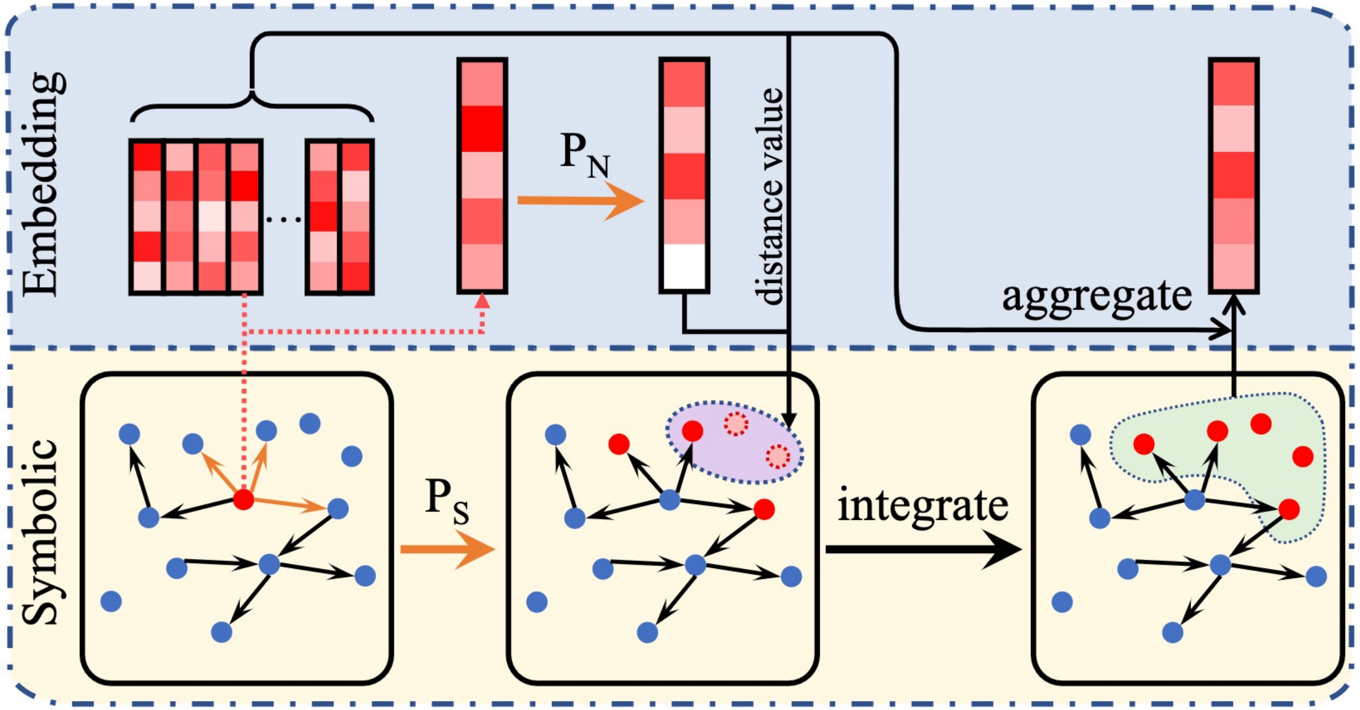


Figure 2: ENeSy's logical operators and the details about neural symbolic entanglement. P_N means neural projection and P_S means symbolic projection.

Neural and Symbolic Entangled Projection Operator



Projection(∃)

$$\mathbf{v}_t = \mathbf{v}_h \circ \mathbf{v}_r, \text{where} |\mathbf{v}_r| = 1$$

$$\mathbf{p}_t = \mathbf{g}(\mathbf{p}_h \mathbf{M}_r)^\top$$

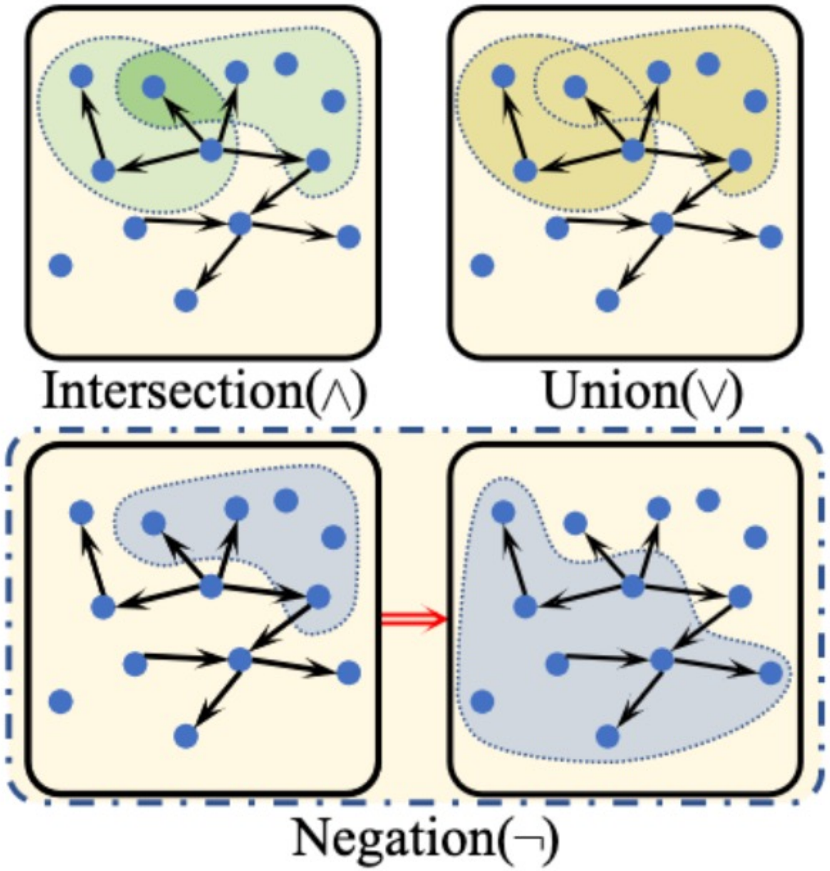
$$\mathbf{S}(\mathbf{x}, \mathbf{y}) = \gamma - \|\mathbf{x} - \mathbf{y}\|_1$$

$$\mathbf{p}''_t = \mathbf{g}(\mathbf{p}_t + \mathbf{p}'_t)$$

$$\mathbf{v}'_t = \sum_{i=1}^{|\mathcal{S}_t|} \mathbf{p}_t^{i\prime\prime} \mathrm{MLP}(\mathbf{v}_{e_i}) \mathbf{v}_{e_i}, e_i \in \mathcal{S}_t$$

Method

Other Logical Operators



First step:

intersection: $\mathbf{p}_1 \wedge \mathbf{p}_2 : \mathbf{g}(\mathbf{p}_1 \circ \mathbf{p}_2),$

union: $\mathbf{p}_1 \vee \mathbf{p}_2 : \mathbf{g}(\mathbf{p}_1 + \mathbf{p}_2 - \mathbf{p}_1 \circ \mathbf{p}_2),$

negation: $\neg\mathbf{p} : \mathbf{g}(\frac{\alpha}{|\mathcal{V}|} - \mathbf{p})$

Second step:

Turn the symbolic results to embedding like the projection operator

$$\mathbf{v}'_t = \sum_{i=1}^{|\mathcal{S}_t|} \mathbf{p}_t^{i\prime\prime}\mathrm{MLP}(\mathbf{v}_{e_i})\mathbf{v}_{e_i}, e_i \in \mathcal{S}_t$$

Method

Neural and Symbolic Ensemble Answering

$$\mathbf{a} = \lambda \mathbf{p} + (1-\lambda)\,\mathrm{Softmax}(\underset{\forall e \in \mathcal{V}}{\mathbf{Concat}}(\mathbf{S}(\mathbf{v}, \mathbf{v}_e)))$$

Loss Function

$$L_1 = -\log\sigma(-\mathbf{S}(\mathbf{v}_q, \mathbf{v}_e)) - \frac{1}{n}\sum_{i=1}^{n}\log\sigma(\mathbf{S}(\mathbf{v}_q, \mathbf{v}_{e'}))$$

$$L_2 = -\log\sigma(\mathbf{p}_e \cdot \log[\mathbf{p}_q^{\top}, \theta]_+)$$

$$L_3 = -\log\sigma(-\mathbf{S}(\mathbf{v}_t, \sum_{i=1}^{|\mathcal{S}_t'|}\mathbf{p}_t^{i\prime}\mathrm{MLP}(\mathbf{v}_{e_i})\mathbf{v}_{e_i})), e_i \in \mathcal{S}_t'$$

Learning Procedure

- **Train with link prediction**

  Train the embedding of entities and relations. **Loss function = L_1**

  Train the MLP function. **Loss function = $L_1 + L_2 + L_3$**

- **Train with complex query**

  Based on the model train ed with only link prediction, use the complex query to fine-tune the model. **Loss function = $L_1 + L_2$**

Dataset & Baselines

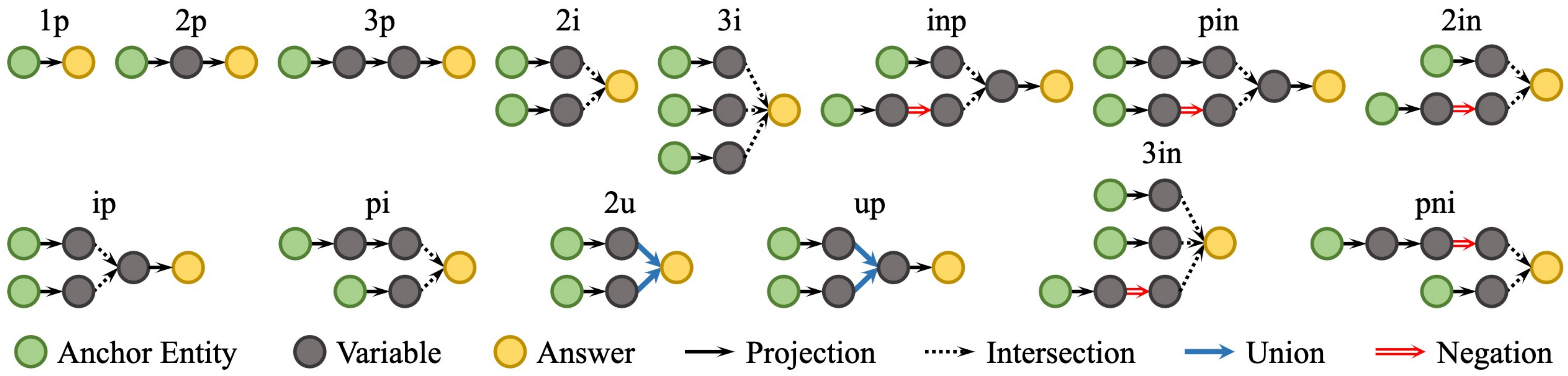


Figure 3: The query structure of all queries used for training and evaluation. Namely, the *p*, *i*, *u* and *n* stands for the projection, intersection, union and negation, respectively.

Experiment

Table 1: The MRR results of FOL queries on FB15K-237 and NELL-995, and the models are trained with only link prediction task. The Avg_p and Avg_n are the average MRR of Existential Positive First Order (EPFO) queries (query with $\exists$, $\vee$ or $\wedge$ and without $\neg$) and queries with $\neg$, respectively. N/A means not available.

Model	Avg_p	Avg_n	1p	2p	3p	2i	3i	pi	ip	2u	up	2in	3in	inp	pin	pni
FB15K-237																
GQE	17.7	N/A	41.6	7.9	5.4	25.0	33.6	16.3	10.9	11.9	6.2	N/A	N/A	N/A	N/A	N/A
Q2B	18.2	N/A	42.6	6.9	4.7	27.3	36.8	17.5	11.1	11.7	5.5	N/A	N/A	N/A	N/A	N/A
BetaE	15.8	0.5	37.7	5.6	4.4	23.3	34.5	15.1	7.8	9.5	4.5	0.1	1.1	0.8	0.1	0.2
FuzzQE	21.8	6.6	44.0	**10.8**	**8.6**	32.3	41.4	22.7	15.1	**13.5**	**8.7**	7.7	9.5	7.0	4.1	4.7
ENeSy	**23.4**	**8.1**	**44.5**	**10.8**	7.7	**33.2**	**48.4**	**25.8**	**18.8**	13.4	7.6	**9.6**	**10.2**	**7.1**	**5.8**	**7.8**
NELL-995																
GQE	21.7	N/A	47.2	12.7	9.3	30.6	37.0	20.6	16.1	12.6	9.6	N/A	N/A	N/A	N/A	N/A
Q2B	21.6	N/A	47.6	12.5	8.7	30.7	36.5	20.5	16.0	12.7	9.6	N/A	N/A	N/A	N/A	N/A
BetaE	19.0	0.4	53.1	6.0	3.9	32.0	37.7	15.8	8.5	10.1	3.5	0.1	1.4	0.1	0.1	0.1
FuzzQE	27.1	7.3	57.6	17.2	**13.3**	38.2	41.5	27.0	19.4	**16.9**	**12.7**	9.1	**8.3**	8.9	4.4	5.6
ENeSy	**28.7**	**9.4**	**58.8**	**17.4**	12.8	**39.1**	**48.9**	**29.1**	**24.1**	16.0	12.4	**10.9**	8.2	**11.0**	**8.4**	**8.6**

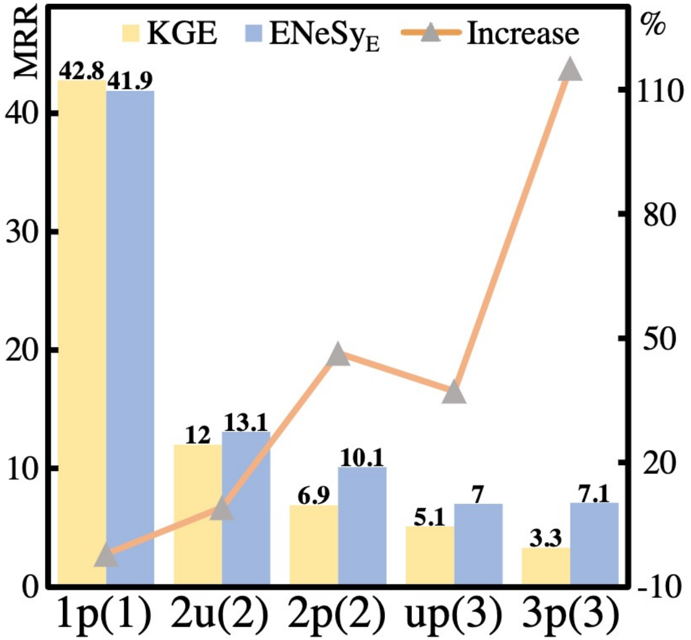
Experiment

Table 2: The average MRR results of FOL queries on FB15K-237 and NELL-995 , and the models are trained with complex query data.

Model	Avg_p	Avg_n	1p	2p	3p	2i	3i	pi	ip	2u	up	2in	3in	inp	pin	pni
FB15K-237																
GQE	16.3	N/A	35.0	7.2	5.3	23.3	34.6	16.5	10.7	8.2	5.7	N/A	N/A	N/A	N/A	N/A
Q2B	20.1	N/A	40.6	9.4	6.8	29.5	42.3	21.2	12.6	11.3	7.6	N/A	N/A	N/A	N/A	N/A
BetaE	20.9	5.5	39.0	10.9	10.0	28.8	42.5	22.4	12.6	12.4	9.7	5.1	7.9	7.4	3.5	3.4
CQD	21.7	N/A	**46.3**	9.9	5.9	31.7	41.3	21.8	15.8	14.2	8.6	N/A	N/A	N/A	N/A	N/A
FuzzQE	24.2	**8.5**	42.2	**13.3**	**10.2**	33.0	47.3	26.2	18.9	**15.6**	**10.8**	9.7	**12.6**	**7.8**	5.8	6.6
ENeSy	**24.5**	**8.5**	44.7	11.7	8.6	**34.8**	**50.4**	**27.6**	**19.7**	14.2	8.4	**10.1**	10.4	7.6	**6.1**	**8.1**
NELL-995																
GQE	18.6	N/A	32.8	11.9	9.6	27.5	35.2	18.4	14.4	8.5	8.8	N/A	N/A	N/A	N/A	N/A
Q2B	22.9	N/A	42.2	14.0	11.2	33.3	44.5	22.4	16.8	11.3	10.3	N/A	N/A	N/A	N/A	N/A
BetaE	24.6	5.9	53.0	13.0	11.4	37.6	47.5	24.1	14.3	12.2	8.5	5.1	7.8	10.0	3.1	3.5
CQD	28.4	N/A	**60.0**	16.5	10.4	**40.4**	49.6	28.6	20.8	16.8	12.6	N/A	N/A	N/A	N/A	N/A
FuzzQE	29.3	8.0	58.1	**19.3**	**15.7**	39.8	**50.3**	28.1	21.8	**17.3**	**13.7**	8.3	**10.2**	11.5	4.6	5.4
ENeSy	**29.4**	**9.8**	59.0	18.0	14.0	39.6	49.8	**29.8**	**24.8**	16.4	13.1	**11.3**	8.5	**11.6**	**8.6**	**8.8**

Model Analysis

Q1: Do symbolic results assist neural reasoning in cascading error?



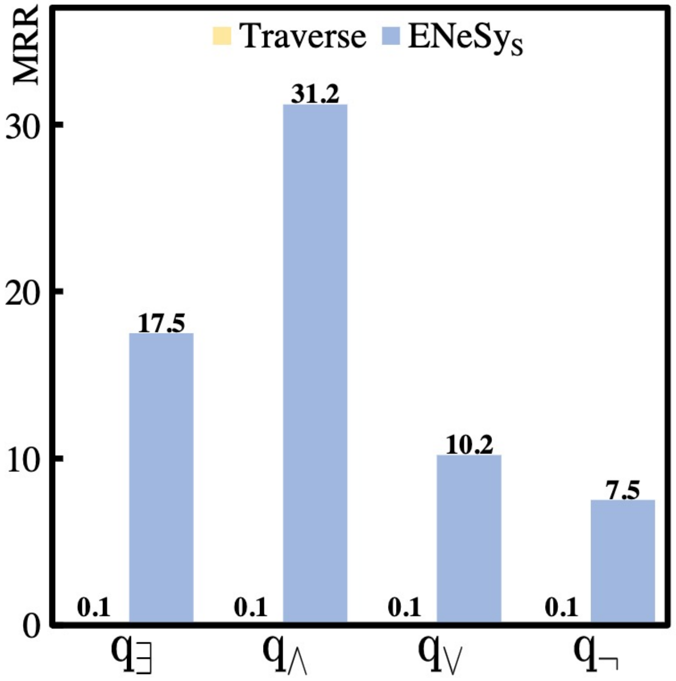
(a)

We compare the pure KE embedding model, and the query types are listed below the horizontal axis and we sort them by the length of the query which is the longest distance from the anchor nodes to the target node in the computation graph.

The MRR results of more complex queries significantly improve with query length increases. This demonstrates that the cascading error, which is the main limitation of multi-hop embedding reasoning, has been alleviated with the symbolic assistant.

Model Analysis

Q2: Does embedding results assist symbolic reasoning in KG incompleteness?



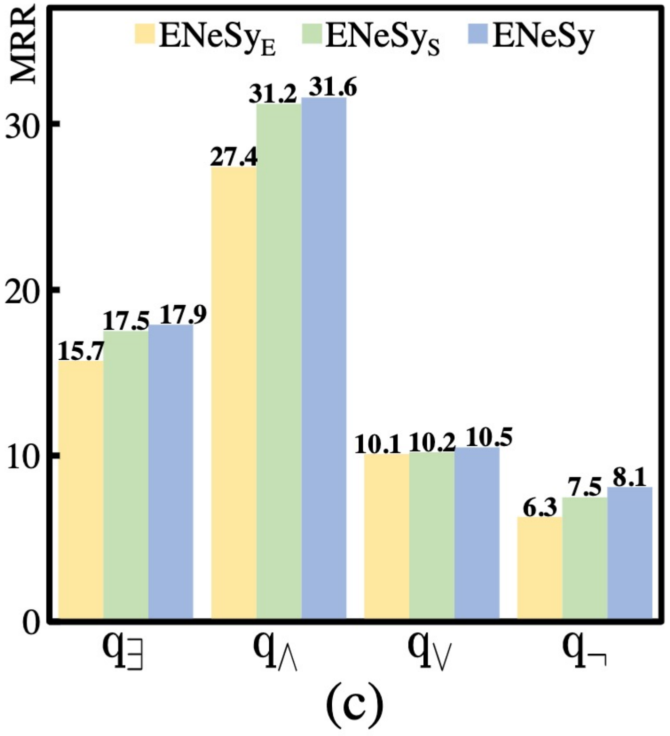
(b)

Since we only evaluate the generalization ability of models with answers that could not be found by simply traversing KG, the traversing results are nearly zero (since the result is MRR, the number won't be an absolute zero), while ENeSy achieves better results than most baselines.

The reason for this significant improvement from zero to almost SOTA performance is in the entangled process.

Model Analysis

Q3: Is ensemble prediction of neural and symbolic results useful?



Ensemble prediction enables us to fuse the symbolic and reasoning results. As the figure illustrates, all the results of different group queries improve with ensemble.

Thank you!

Our code and data are available at: https://github.com/zjukg/ENeSy👉

Contact information: xuzezhong@zju.edu.cn, zhang.wen@zju.edu.cn