

# Supplemental Material for "A Holistic Approach for Answering Logical Queries on Knowledge Graphs"

Yuhan Wu

East China Normal University  
yuhanwu0001@163.com

Xuemin Lin

Shanghai Jiao Tong University  
lxue@cse.unsw.edu.au

Yuanyuan Xu

Zhejiang Lab, University of New South Wales  
xyycn8@163.com

Wenjie Zhang

University of New South Wales  
wenjie.zhang@unsw.edu.au

## I. THE PROPOSED HALK

TABLE I: Notations and definitions.

Symbols	Definition
$\mathcal{G} = \{V, R, T\}$	the knowledge graph
$v_i$	the $i$ -th entity/node in knowledge graph
$\mathbf{v}_i$	the embedding of entity/node $v_i$
$r_i$	the $i$ -th relation/edge in knowledge graph
$\mathbf{e}_i$	the embedding of relation/edge $r_i$
$\mathcal{Q} = \{U, R, L\}$	a dependency graph
$U$	the node set of $\mathcal{Q}$
$u_i$	the $i$ -th node in $U$
$L$	the set of Logical operations
$\mathbf{A}_i$	the arc embedding of $U_i$
$\mathbf{A}_{i,c}$	the center points of $\mathbf{A}_i$
$\mathbf{A}_{i,l}$	the arclengths of $\mathbf{A}_i$
$\mathbf{A}_{i,S}$	the start points of $\mathbf{A}_i$
$\mathbf{A}_{i,E}$	the end points of $\mathbf{A}_i$

### A. Semantic Center

We illustrate it using an example in Fig. 1. Without loss of generality, let  $d = 1$ , when  $\mathbf{A}_{1,c} = (\rho, \alpha)$  and  $\mathbf{A}_{2,c} = (\rho, 3\pi/2 + \beta)$  ( $0 < \alpha, \beta < \pi/2$ ), the common weighted average point of  $\mathbf{A}_{1,c}$  and  $\mathbf{A}_{2,c}$  may lead to wrong average center point that is closed to  $(\rho, \pi)$  (see Fig. 1a), yet, actually, the expected semantic center point should be around  $(\rho, 0)$  (see Fig. 1b). In a word, due to the periodicity of angle, the usual weighted average may lead to inconsistent semantics.

### B. Difference operation

*Lemma 1:* The boundary of answer region returned by HaLk is tighter than the one returned by NewLook.

*Proof 1:* In fact, as shown in Fig. 3a, Newlook [1] would generate false negatives in  $\mathbf{a}_2$  and  $\mathbf{a}_3$  when only modelling answer area  $\mathbf{a}_1$ . If more answers are required, it would include false positives by expanding the box area to the yellow dashed rectangle, that is, covering the unexpected entities in  $\mathbf{b}_2$  and  $\mathbf{b}_3$ . On the other hand, HaLk is able to completely model the

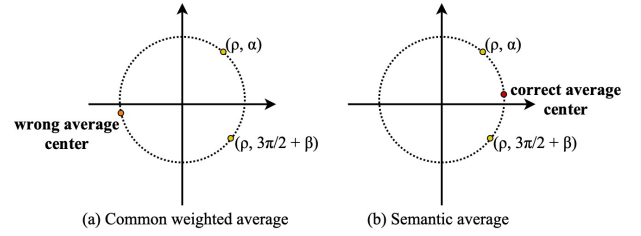


Fig. 1: Illustrations of two kinds of center average protocols.

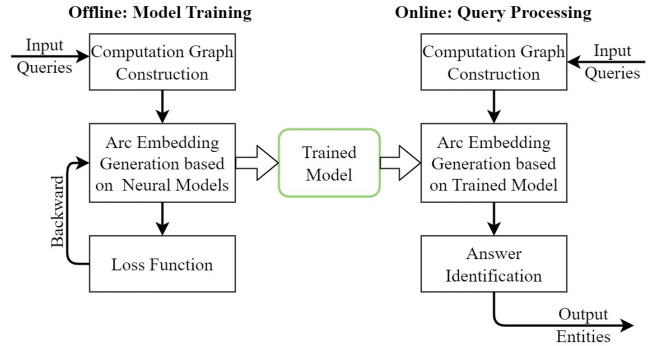


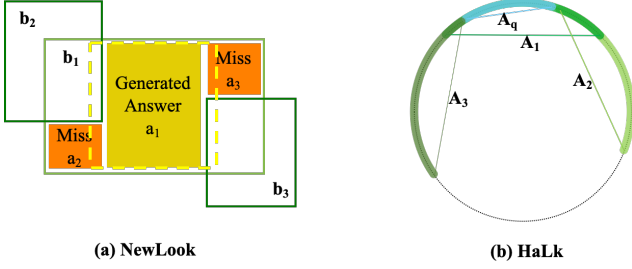
Fig. 2: Framework of our proposed HaLk.

answer area by arcs as shown in Fig. 3b since the answer area  $\mathbf{A}_q$  of  $\mathbf{A}_1 - \mathbf{A}_2 - \mathbf{A}_3$  is still a closure of arc. Thus, the theorem holds.

*Remarks.* Here are two special situations, that is,  $\mathbf{b}_1/\mathbf{A}_1$  is entirely inside  $\mathbf{b}_i/\mathbf{A}_i$ , or multiple boxes/arcs have no overlap. Both NewLook and HaLk can completely model answer regions via attention neural network, yet these two situations are rarely raised in real-world scenarios. For example, "What plants are living in the desert but not living under water?"

### C. Framework of HaLk

Fig. 2 shows the overall framework of the proposed HaLk including the offline stage and online stage. During the training stage, for the input training query structures, we need to first construct the corresponding computation graph, and then



$c$  and target node  $d$  both hit the randomly selected ground truth.

Fig. 3: Illustrations of two models for difference operator.

traverse the computation graph starting from anchor node(s) to iteratively solve the sub-queries til get the arc embedding of the target node of the query. After that, we calculate the loss and make a gradient back propagation [2] to update the embeddings of the entities, the embeddings of the relations, and the parameters of all the logical operators' networks. The training will continue until the predetermined maximum iterative steps or model convergence is reached. As for online query stage, the trained model would generate the arc embedding for the target node of the input test logical query by executing a set of logical operations. According to the obtained arc embedding, we measure the similarity between entities and the query by the distance between them in the vector space. At last, all the entities that are inside or close to the query's arc embedding would be selected as answers to the query.

## II. EXPERIMENTS

### A. Training and evaluation query structures.

As shown in Fig. 4, the query structures on the top line are considered to verify HaLk's ability to model the newly-defined FOL operations, (i.e., projection, intersection, difference and union). And the complex query structures including 'ip', 'pi', '2u', 'up' and 'dp' would only be evaluated on validation and test stage to test the generalization performance of HaLk. The additional query structures with negation operator on the bottom line are used to evaluate model's performance on queries with negation operation together with the top line training query structures.

### B. Visualization of Answering Logical Queries

To explain more clearly about how to answer a logical query on knowledge base on HaLk, we visualize the intermediate answer entities and target answer entities for a given logical query that was generated from a SPARQL [3] query to answer the question "What company was founded by a student from West Point and was also the one that Buzz Aldrin work for?" on FB15k-237 [4] dataset. Since the intermediate results for the variables are also arc embeddings, we can also get the prediction results for the intermediate variables via LSH in the embedding space.

Combined with the variable nodes marked in Fig. 5, we can see from Table II that although the prediction results of intermediate variable nodes  $a$  and  $b$  do not hit the randomly selected positive entity, the prediction results of variable node

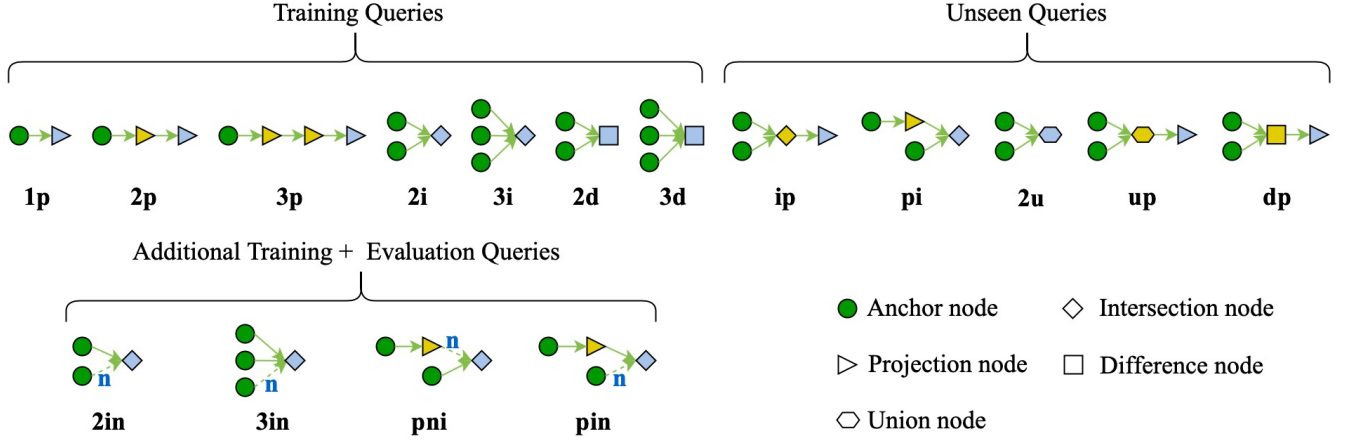


Fig. 4: Query structures with abbreviation of their computation graph used in the experiments, where 'p', 'i', 'd', 'n' and 'u' represent 'projection', 'intersection', 'difference', 'negation' and 'union'.

TABLE II: Visualization of a pi query on FB15k-237 test set.

Query	$q = d_? \cdot \exists a : Student(West\ Point, a) \wedge Found(a, d_?) \wedge Company(Buzz\ Aldrin, d_?)$		
Variable	Top 6 Prediction Results		Positive Sample
a	David Randolph Scott Simone Askew Roy Moore	Roger Hilsman Jr. Frank Frederick Borman II Mark Thomas Esper	Donald Herod Peterson
b	NSA NASA Scrum	Colgate-Palmolive CIA US Department of Housing and Urban Development	US Department of the Air Force
c	US Air Force Academy Microsoft RAND Corporation	The Walt Disney Company NASA Omega SA	NASA
d	Apple NASA CBS	Pixar Animation Studio US Air Force Academy Slitherine Software	NASA

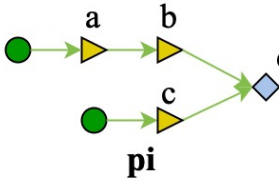


Fig. 5: Computation graph of question "What company was founded by a student from West Point and was also the one that Buzz Aldrin work for?".

## REFERENCES

- [1] L. Liu, B. Du, H. Ji, C. Zhai, and H. Tong, "Neural-answering logical queries on knowledge graphs," in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. ACM, 2021, pp. 1087–1097.
- [2] S.-i. Amari, "Backpropagation and stochastic gradient descent method," *Neurocomputing*, vol. 5, no. 4-5, pp. 185–196, 1993.

- [3] L. Zou, J. Mo, L. Chen, M. T. Özsu, and D. Zhao, "gstore: answering sparql queries via subgraph matching," *Proceedings of the VLDB Endowment*, vol. 4, no. 8, pp. 482–493, 2011.
- [4] K. D. Bollacker, C. Evans, P. K. Paritosh, T. Sturge, and J. Taylor, "Freebase: a collaboratively created graph database for structuring human knowledge," in *Proceedings of the ACM SIGMOD International Conference on Management of Data*. ACM, 2008, pp. 1247–1250.