Improving Knowledge Base Question Answering with Question Understanding Augment

Peiyun Wu and Xiaowang Zhang

College of Intelligence and Computing, Tianjin University, Tianjin 300350, China
{wupeiyun,xiaowangzhang}@tju.edu.cn **

Abstract. The basement of knowledge base question answering (KBQA) is to understand the given question and extract the meaning from it. Existing works largely focus on generating query graphs to represent the semantics of the question while ignoring understand the real meaning of it. To augment question understanding, in this paper, we leverage rich external linguistic knowledge to enhance question semantics. First, we integrate the sememe and gloss information into words, where sememe (the minimum semantic units of word meanings) and gloss (sense definition) are used to disambiguate the word sense and enrich questions information. Moreover, we present a co-attention network to build co-dependent representations for the sememe and gloss. Experiments evaluated on two data sets show that our model outperforms existing approaches.

1 Introduction

Semantic parsing is an important approach to KBQA, which constructs a query structure (called *query graph*) that represents the semantics of questions. Semantic parsing based approaches effectively transform questions into logical forms where the reliability of logical forms can ensure the correctness of answering questions. The success of semantics parsing lies in representing the semantics of questions to better capture users' intention.

However, in recent years, many semantic parsing approaches focus on complex query graph generation and re-ranking[1, 2, 5] with paying little attention to understanding the meanings of questions accurately. They aim to leverage a ranking model to score and find the best query graph. As a result, existing works without processing ambiguous questions cannot always perform query graph ranking better.

In this paper, we propose an augmented question representation method by leveraging the sememe and gloss information. Specifically, we take advantage of the gloss information from WordNet and the sememe [3] information from Hownet into word embeddings from the given question. A word may have multiple senses and a sense consists of several sememes and a gloss. To highlight important information in the sememe and gloss, we present a co-attention network to generate better representations.
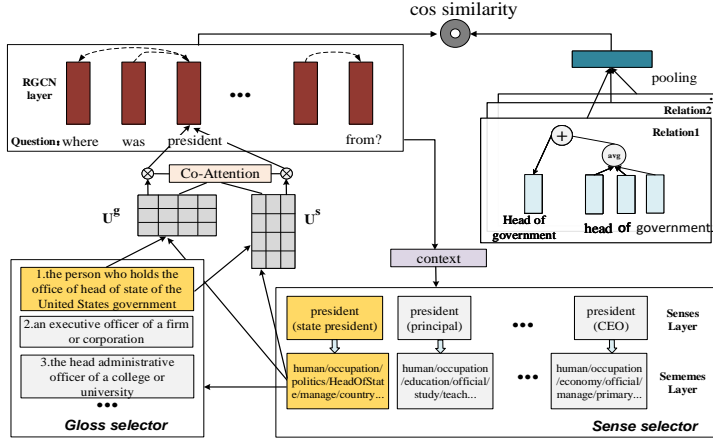
Fig. 1: Diagram for Our Model.

2 Our Approach

Given a question $Q = \{w_1, \ldots, w_n\}$, we generate its candidate query graph set by method in [1]. We measure the semantic similarities between the question and each query graph to find the optimal one. For word w_i in Q, we denote the set of its senses as S^{w_i}, and $E^{w_i} = \{e_{i1}, \ldots, e_{ik}\}$ represents an unordered set of all sememes contained in w_i. We assume for word w_i that we have a gloss set G^{w_i}. Our model is shown in Fig. 1.

Sense Selector: This selector selects each word sense that is most relevant to the context. We first feed Q into a bi-directional long short-term memory network (Bi-LSTM) to generate the hidden representation $\{h_1, \ldots, h_n\}$. Then we generate the context representation of Q as below:

$$context = \sum_{i=1}^{n} \text{softmax}(tanh(h_i^\top \cdot (\frac{1}{n}\sum_{i=1}^{n} h_i))) \cdot h_i \qquad (1)$$

To calculate the correlation between each sememe and context, we use Sigmoid function to obtain probability value by:

$$p(e_{ij}|context) = \text{Sigmoid}\left(context \cdot e_{ij}^\top\right), \quad \forall\, j \in (1, \ldots, k) \qquad (2)$$

For each sense in S^{w_i}, its probability is calculated by the average of all the sememe probabilities that it contains. In this way, we can select the sense with the highest probability value under the current context and denote it as $S_{max}^{w_i} = \{e_1, \ldots, e_k\}$, where $\{e_1, \ldots, e_k\}$ represents all sememe vectors it contains.

Gloss Selector: This selector selects the gloss that is most relevant to the selected word sense. Analogously, we use Sigmoid function to obtain the highest probability of gloss that is most relevant to the average embedding of $S_{max}^{w_i}$. We denote the selected gloss as $G_{max}^{w_i} = \{o_1, \ldots, o_m\}$, where $\{o_1, \ldots, o_m\}$ represents all word embeddings it contains.

Co-Attention: To model the mutual influence and highlight the important information in the sememe and gloss, we introduce a co-attention network to dynamically combine the sememe and gloss representation as:

$$U^s = \tanh(S_{max}^{w_i} \cdot {G_{max}^{w_i}}^\top), \quad U^g = \tanh(G_{max}^{w_i} \cdot {S_{max}^{w_i}}^\top) \tag{3}$$

$$SG^{w_i} = \lambda \sum_{j=1}^{k} [U^s \cdot \mathrm{softmax}(U_{:}^{s})]_{:j} + (1-\lambda) \sum_{j=1}^{m} [U^g \cdot \mathrm{softmax}(U_{:}^{g})]_{:j} \tag{4}$$

where SG^{w_i} is the combination representation of the selected sememe and gloss of word w_i, λ is the parameter and $\lambda \in [0,1]$. $\mathrm{softmax}(U_{:}^{s})$ and $\mathrm{softmax}(U_{:}^{g})$ are attention weight matrix for softmax function across each column of U^s and U^g, respectively. $[]_{:j}$ denotes j-th column of $[]$. Finally, we concatenate SG^{w_i} to the word embedding of w_i to enrich the semantics and reduce ambiguity.

Question Representation: We treat $\{x_1, \ldots, x_n\}$ as the initial representations of the given question which has been integrated with sememe and gloss. To further augment the contextual embeddings of the question, we parse the question into its syntactic dependencies graph D_G and adopt relational graph convolutional network (RGCN) to digest this structural information:

$$x_i^{(l+1)} = \mathrm{ReLU} \left(\sum_{r \in R} \sum_{j \in N_i^r} \frac{1}{|N_i^r|} W_r^{(l)} x_j^{(l)} + W_0^{(l)} x_i^{(l)} \right) \tag{5}$$

Here R is a set of dependency-relation. l is l-th layer, N_i^r is the set of all r-neighbors of i-th node in D_G. Note that W_0 and W_r are weighted matrixes. Finally, we apply a pooling operation after the last RGCN layer to get the representation of the question.

Relation Representation We represent relations from different granularity in a query graph. For each relation, we take its relation-level and word-level representations into consideration. The word-level relation is calculated by its average word embeddings. The relation-level representation is the vector of unique token of relation name. Then, each relation is represented by the sum operation and we perform max pooling over relations to obtain the final relation representation.

3 Experiments and Evaluations

We use Wikidata as our KB and conduct on two data sets, namely, *WebQSP-WD*(WSPWD)[1] and *QALD-7* (Task 4, English), both support for Wikidata. We use F1-score as metrics, where all results are macro-averaged scores.

By Table 1, we show that our model outperforms all datasets. Our model achieves 54.2%, 23.3%, 8.9%, 11.9% higher F1-score compared to STAGG, HR-BiLSTM, GGNN, Slot-Matching on WSPWD. Analogously, we achieve 59.3%, 45.7%, 39.1%,21.7% higher F1-score on QALD-7. We observe that if we only integrate sememe information "+sememe" or gloss information "+gloss", our model performs worse but still keep competitive. We can conclude the effectiveness of our augmented question representation with sememe and gloss integration.
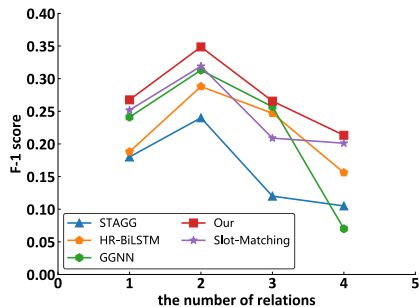
Fig. 2: The number of relations to find correct answers

Table 1: Overall Average Results over Wikidata

Model	WSPWD	QALD-7
STAGG(2015) [2]	0.1828	0.1861
HR–BiLSTM(2017) [5]	0.2287	0.2035
GGNN(2018)[1]	0.2588	0.2131
Slot-Matching(2019)[4]	0.2519	0.2436
+sememe	0.2459	0.2546
+gloss	0.2597	0.2743
Our	**0.2819**	**0.2965**

To measure the performance across questions of different complexity, we break down the performance by the number of relations that are needed to find the correct answer on WSPWD, and results are shown in Fig. 2, we can see that our model is effective in dealing with different complexity questions.

4 Conclusion

In this paper, we augment question understanding to improve KBQA. The sememe and gloss information can benefit from each other and enhance question semantics. In this way, our approach provides a new method of usage of external knowledge in question representation. In future work, we are interested in extending our model for more complex practical questions.

5 Acknowledgments

References

1. Sorokin, D., Gurevych, I.: Modeling semantics with gated graph neural networks for knowledge base question answering. In: *COLING'2018*, pp.3306–3317.
2. Yih, W., Chang, M., He, X., Gao, J.: Semantic parsing via staged query graph generation: question answering with knowledge base. In: *ACL'2015*, pp.1321–1331.
3. Bloomfield, L.: A set of postulates for the science of language. Language, 1926, **2**(3): 153-164.
4. Maheshwari, G., Trivedi, P., Lukovnikov, D., Chakraborty, N., Fischer, A., Lehmann, J. (2019). Learning to rank query graphs for complex question answering over knowledge graphs. In: ISWC'19, pp.487-504.
5. Yu, M., Yin, W., Hasan, K.S., Santos, C.N., Xiang,B., Zhou,B.: Improved neural relation detection for knowledge base question answering. In: *ACL'2017*, pp.571–581.